

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften: Beiträge zu methodischen Problemen des ALLBUS 1980

Mayer, Karl Ulrich (Ed.); Schmidt, Peter (Ed.)

Veröffentlichungsversion / Published Version

Sammelwerk / collection

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Mayer, K. U., & Schmidt, P. (Hrsg.). (1984). *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften: Beiträge zu methodischen Problemen des ALLBUS 1980* (Monographien: Sozialwissenschaftliche Methoden, 5). Frankfurt am Main: Campus Verlag. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-75326-7>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Karl Ulrich Mayer,
Peter Schmidt (Hg.)

Allgemeine Bevölkerungs- umfrage der Sozialwissenschaften

Zentrum für Umfragen, Methoden
und Analysen (ZUMA), Mannheim
Monographien Sozialwissen-
schaftliche Methoden

Campus

Zentrum für Umfragen, Methoden und Analysen
(ZUMA), Mannheim
Monographien Sozialwissenschaftliche Methoden
Band 5

Herausgegeben von
Max Kaase und Manfred Küchler

Wissenschaftlicher Beirat

Hubert Feger, Walter Jaide, M. Rainer Lepsius, Franz Urban
Pappi, Hansgert Peisert, Erwin K. Scheuch, Werner Tack,
Rudolf Wildenmann, Wolfgang Zapf und Rolf Ziegler

Karl Ulrich Mayer, Peter Schmidt (Hg.)

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften

Beiträge zu methodischen Problemen
des ALLBUS 1980

Campus Verlag
Frankfurt/New York

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften

: Beitr. zu method. Problemen d. ALLBUS 1980 / Karl
Ulrich Mayer ; Peter Schmidt (Hg.). – Frankfurt/Main ;
New York : Campus Verlag, 1984.

(Monographien : Sozialwiss. Methoden ; Bd. 5)

ISBN 3-593-33262-0

NE: Mayer, Karl Ulrich [Hrsg.]; Monographien /
Sozialwissenschaftliche Methoden

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung
sowie der Übersetzung, vorbehalten. Kein Teil des Werkes darf in irgendeiner
Form (durch Photokopie, Mikrofilm oder ein anderes Verfahren) ohne
schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung
elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Copyright © 1984 Campus Verlag GmbH, Frankfurt/Main

Umschlaggestaltung: Eckard Warminski, Frankfurt/Main

Druck und Bindung: Beltz Offsetdruck, Hemsbach

Printed in Germany

I n h a l t

Vorwort	7
Karl Ulrich Mayer Zur Einführung: Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften als eine Mehrthemen-Wiederholungsbefragung	11
Hartmut Esser Determinanten des Interviewer- und Befragtenverhaltens: Probleme der theoretischen Erklärung und empirischen Untersuchung von Interviewereffekten	26
Volker Schanz und Peter Schmidt Interviewsituation, Interviewermerkmale und Reaktionen von Befragten im Interview: eine multivariate Analyse	72
Hans-Peter Kirschner ALLBUS 1980: Stichprobenplan und Gewichtung	114
Jürgen H.P. Hoffmeyer-Zlotnik Erfassen von Wohnquartiersvariablen - ein Mittel zur soziologischen Zuordnung der Wohnbevölkerung	183
Manfred Küchler Eine sozio-demographische Beschreibung der Träger post-materialistischer Einstellungen	215
Cornelia Krauth und Rolf Porst Sozioökonomische Determinanten von Einstellungen zu Gastarbeitern	233
Peter Schmidt und Gunter Wolf Sozialstrukturelle und individuelle Determinanten von subjektiver Schichtidentifikation und politischen Einstellungen	267
Andreas Diekmann Einkommensdiskriminierung von Frauen - Messung, Analyseverfahren und empirische Anwendungen auf Angestellteinkommen in der Bundesrepublik	315

V o r w o r t

Der hier vorgelegte Band enthält Beiträge zu methodischen Problemen der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) auf der Grundlage der im Frühjahr 1980 durchgeführten Repräsentativbefragung.

Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ist ein langfristig angelegtes Forschungsprogramm, das sowohl für Lehre und Forschung aktuelle Repräsentativdaten über demographische Merkmale, Einstellungen und Verhaltensweisen zur Verfügung stellen als auch mit Hilfe von Wiederholungsbefragungen der gesellschaftlichen Dauerbeobachtung und der Analyse des sozialen Wandels dienen soll.

Die Beiträge in diesem Band verdanken ihre Entstehung der zusätzlichen Zielsetzung, mit dem ALLBUS die Methodik und Praxis der sozialwissenschaftlichen Umfrageforschung weiterzuentwickeln. Wir verbinden damit die Absicht, den Benutzern der ALLBUS-Daten sowohl deren methodische Grundlagen und Probleme sichtbar zu machen als auch die vielfältigen Möglichkeiten bei der Auswertung dieser Daten aufzuzeigen. Die Beiträge sind jedoch über den Kreis der ALLBUS-Benutzer hinaus für solche Leser geschrieben worden, die an methodischen Problemen der Umfrageforschung und an empirischen Analysen über die Bevölkerung der Bundesrepublik interessiert sind. Sie richten sich damit insbesondere auch an Studenten in Kursen über Methoden der empirischen Sozialforschung.

Unter systematischen Gesichtspunkten fallen die Arbeiten in zwei Gruppen. Esser, Schanz/Schmidt, Kirschner und Hoffmeyer-Zlotnik befassen sich mit methodischen Problemen, die die Umfrageforschung allgemein betreffen. Während Schanz und Schmidt ein empirisches Modell der Bedingungen von Interviewereinflüssen entwickeln und an den ALLBUS-Daten prüfen, entfaltet Esser eine sozialpsychologische Theorie der wechselseitigen Einflüsse in der Interviewsituation und der Richtung ihrer Wirkungen. Kirschner behandelt ein weitgehend vernachlässigtes Problem der Methodik der Umfrageforschung, nämlich die stichproben-theoretischen Voraussetzungen und Möglichkeiten, Ausfälle (z.B. durch Verweigerung oder Nichtantreffbarkeit) durch geeignete Gewichtungungsverfahren zu korrigieren. Das Ausfallproblem war auch der Anlaß der Untersuchung von

Hoffmeyer-Zlotnik. Er berichtet über die Entwicklung eines Verfahrens zur Messung sozial-morphologischer Charakteristika von Wohnquartieren, das Aussagen über die Struktur von Ausfällen bei Befragungen erlauben soll.

Die zweite Gruppe von Arbeiten setzt sich aus inhaltlichen Analysen mit den Daten der als Mehrthemenbefragung konzipierten ALLBUS-Studie zusammen. Die Auswahl gerade dieser Untersuchungen soll veranschaulichen, mit welchen unterschiedlichen methodischen Zielsetzungen und Mitteln der Datenanalyse ALLBUS-Informationen ausgewertet werden können. Küchler verfolgt mit der Frage nach dem Erklärungswert sozialer Hintergrundmerkmale postmaterialistischer Einstellungen einen "klassischen" Untersuchungsansatz der Umfrageforschung, wendet dabei jedoch ein relativ neues Verfahren der multivariaten Analyse nicht-metrisch gemessener Daten an, das mit der Zerlegung von Prozentdifferenzen ein besonders anschauliches Instrument der Sozialforschung bietet. Krauth und Porst verwenden in ihrer Studie über Determinanten der Einstellungen zu Gastarbeitern komplexe Strukturgleichungsmodelle mit einem voraussetzungsreichen Meßniveau. Schmidt und Wolf untersuchen den Einfluß des sozialen Kontextes auf die subjektive Schichtidentifikation und politische Einstellungen und demonstrieren, wie Kontexteffekte mit einem Strukturgleichungsmodell geprüft werden können. Diekmanns Studie über Einkommensdiskriminierung bei weiblichen Erwerbstätigen kann als ein gutes Beispiel für Sozialstrukturanalysen mit Umfragedaten dienen, d.h. von Untersuchungen, die häufiger mit Daten der amtlichen Statistik durchgeführt werden und nur "objektive" Informationen erfordern.

Das Zentrum für Umfragen, Methoden und Analysen (ZUMA) e.V. in Mannheim trägt die methodische und organisatorische Verantwortung für die Durchführung der ALLBUS-Umfragen. Für ZUMA betreuten die Herausgeber den ALLBUS als wissenschaftlicher Leiter (1979-1983) und Projektleiter (1980-1981). ZUMA arbeitete dabei im Auftrag und in enger Kooperation mit den Antragstellern und DFG-Mitarbeitern des ALLBUS-Projekts (Prof. M. Rainer Lepsius, Heidelberg; Prof. Erwin K. Scheuch, Köln; Prof. Rolf Ziegler, München; Dipl.-Soz. Rolf Porst und Dipl.-Soz. Cornelia Krauth).

Cornelia Krauth danken wir für ihre engagierte Hilfe bei den Editionsarbeiten, ebenso Irene Meckler für die Erstellung der Schaubilder sowie Marie Louise Frank, Dagmar Haas und Margarete Knörzer für das Schreiben der Manuskripte.

Der Deutschen Forschungsgemeinschaft und dem Bundesministerium für Forschung und Technologie ist zu danken für die Förderung des ALLBUS-Projekts und die Bereitstellung von Mitteln für diese Veröffentlichung.¹

Berlin und Gießen, im August 1983

Karl Ulrich Mayer
Peter Schmidt

¹ Die vorliegende Veröffentlichung wird ergänzt durch zwei weitere Publikationen:

1. Allgemeine Bevölkerungsumfrage der Sozialwissenschaften - ALLBUS 1980, Codebuch mit Methodenbericht und Vergleichsdaten, hrsg. vom Zentralarchiv für empirische Sozialforschung, Köln, und dem Zentrum für Umfragen, Methoden und Analysen (ZUMA) e.V., Mannheim, 1982.
2. Rolf Porst, Allgemeine Bevölkerungsumfrage der Sozialwissenschaften: Ziele, Anlage, Methoden und Resultate (Studienkurs des Fachbereichs Erziehungs- und Sozialwissenschaften der Fernuniversität Hagen), Hagen 1983.

Die Daten und Codebücher der ALLBUS-Umfragen 1980, 1982 und 1984 sind gegen eine geringe Unkostenvergütung zu beziehen über das Zentralarchiv für empirische Sozialforschung, Bachemer Str. 40, 5000 Köln 41.

Anfragen zum ALLBUS-Projekt und weiteren Projektdokumentationen richten Sie bitte an Max Haller oder Rolf Porst, ZUMA, B 2, 1, 6800 Mannheim 1, Tel. 0621/12003.

Karl Ulrich Mayer

Zur Einführung: Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften als eine Mehrthemen-Wiederholungsbefragung

Diese Einführung soll dazu dienen, den Typus der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften als einer Mehrthemen-Wiederholungsbefragung methodisch einzuordnen. Sie setzt an bei der Entstehungsgeschichte dieses Forschungsprogramms und den ursprünglich damit verbundenen Zielsetzungen und berichtet über die Erweiterung dieser Ziele im Verlauf der praktischen Umsetzung der Projektidee. Mit einer Explikation des Kriteriums der Replikation und der spezifischen Leistungen des ALLBUS im Ensemble der sozialwissenschaftlichen Datenerhebungsverfahren soll ferner sichtbar gemacht werden, welcher systematische Stellenwert dem ALLBUS-Forschungsprogramm zukommt.

1. Entstehungsgeschichte und Zielsetzungen

Der Plan eines "Nationalen Surveys" für die Bundesrepublik Deutschland wurde zum erstenmal im Zusammenhang mit der Forschungsplanung der Deutschen Forschungsgemeinschaft, dem sogenannten "Grauen Plan", im Jahre 1975 von Wolfgang Zapf entwickelt. In Anknüpfung an ausländische Vorbilder, wie z.B. dem General Social Survey des National Opinion Research Center in den USA, der skandinavischen "Welfare"- und "Standard of living"-Umfragen und dem englischen General Household Survey, sollte auch in der Bundesrepublik eine "regelmäßige Erhebung von Einstellungen, Zufriedenheiten, Aspirationen und Wertorientierungen sowie der Lebensweisen und Lebensumstände der Bevölkerung" durchgeführt werden.

Im Auftrag der Senatskommission für empirische Sozialforschung der DFG hat eine Gruppe von Sozialwissenschaftlern - darunter maßgeblich Franz Urban Pappi - in den darauffolgenden Jahren eine Konzeption für eine Allgemeine Bevölkerungsumfrage der Sozialwissenschaften erarbeitet.

Dieser Konzeption zufolge sollte in regelmäßigen zeitlichen Abständen ein repräsentativer Querschnitt der Bevölkerung mit einer Zufallsstichprobe

von etwa 3000 Personen mit einem teils stetigen, teil rotierenden Fragenprogramm befragt werden.

Das Fragenprogramm sollte so angelegt sein, daß schon durch die Wiederholung von inhaltlich und methodisch bewährten Fragen aus früheren sozialwissenschaftlichen Umfragen zeitlich und z.T. international vergleichende Analysen möglich und Zeitreihen begründet würden.

Durch die Berücksichtigung verschiedener Themen- und Problembereiche sollte zumindest auf mittlere Sicht eine Art Gesamtbeschreibung und -analyse der Gesellschaft und ihres Wandels ermöglicht werden.

Die beiden Zielsetzungen einer eher "akademisch" orientierten Analyse des sozialen Wandels einerseits und der politikorientierten Sozialberichterstattung bzw. gesellschaftlichen Dauerbeobachtung andererseits belegen schon im Ansatz sich zumindest teilweise widersprechende Optionen für das ALLBUS-Programm. Beschreibungen in der Tradition der Sozialindikatorenbewegung setzen ein möglichst vielfältiges und anschauliches Raster von Dimensionen und Merkmalen voraus und erfordern eine Priorität in der möglichst raschen Gewinnung vieler Meßpunkte. Diese Perspektive legt es im übrigen nahe, bei der Auswahl von Fragen insoweit pragmatisch zu verfahren, als deren deskriptiver Gehalt wichtiger wäre als eine eindeutige theoretische Zurechenbarkeit.

Im Gegensatz dazu setzt eine komplexe, kausal orientierte Analyse des sozialen Wandels im Grunde genommen eine relativ geschlossene Gesellschaftstheorie voraus, die das dynamische Wechselspiel institutioneller Differenzierung, gesellschaftlicher Spaltungen, Bindungen und Wertmilieus empirisch in einigen wenigen, strategisch ausgewählten Merkmalen abzubilden gestatten würde. Dementsprechend erhöhen sich die Begründungsansprüche an die Auswahl von Fragen, deren theoretische Operationalisierungsgüte und systematischen Zwischenbezüge in einem Maße, daß sie mit aus älteren Untersuchungen zur Verfügung stehenden Fragen in der Regel weder methodisch noch inhaltlich voll eingelöst werden können.

Zu den genannten beiden Zielsetzungen kam jedoch von Anfang an eine ganz praktische hinzu. Der ALLBUS sollte nicht nur in dem Sinne "allgemein" sein, daß er als eine Gemeinschaftsaufgabe der Sozialwissenschaften eingerichtet und die Fragenauswahl einen breiten Konsensus unter vielfältigen Beteiligungsmöglichkeiten widerspiegeln sollte. Diese

Daten sollten insbesondere als eine gemeinsame Ressource schon unmittelbar nach der ersten Umfrage jedermann zu geringen Kosten zur Verfügung stehen. Dies bedeutete jedoch zugleich, daß jede einzelne Umfrage auch unbeschadet der längerfristigen Zielsetzungen für eine möglichst große Zahl interessierter Nutzer in Forschung und Lehre analysefähig zu sein hatte.

Diese drei Ziele haben die Fragenprogramme des ALLBUS 1980, 1982 und 1984 in unterschiedlicher Gewichtung bestimmt. Die in diesem Band veröffentlichten Arbeiten beziehen sich ausschließlich auf das zuletzt genannte Kriterium, das Analysepotential der ALLBUS-Umfragen als je einzelnen Querschnittsbefragungen, während die in dem Codebuch dokumentierten Zeitreihen die Möglichkeiten der beiden zuerst aufgeführten Zielsetzungen der gesellschaftlichen Dauerbeobachtung und Analyse des sozialen Wandels erkennen lassen.

In den bisher dreieinhalb Jahren der Durchführung der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften ist jedoch deutlich geworden, daß dieses Forschungsprogramm wichtige zusätzliche Aufgaben erfüllen kann.

Der ALLBUS ist zu einem bedeutsamen Instrument der Methodenentwicklung im Bereich der nicht-kommerziellen Umfrageforschung geworden. Durch die explizite Festlegung und damit Überprüfbarkeit methodischer Standards, eine umfangreiche Dokumentation der Stichproben und Erhebungsverfahren und die bewußte Offenlegung von Stichproben- und Meßfehlern ist der ALLBUS nicht zuletzt als eine exemplarische Bevölkerungsumfrage für die Verwendung in der Lehre der Methoden der empirischen Sozialforschung geeignet. Der vorliegende Band ist ein Teil dieser Bemühungen, die nur auf der Grundlage der Integration des ALLBUS-Projekts in die Dienstleistungsaufgaben von ZUMA möglich waren.

Nicht zuletzt kommen dem ALLBUS jedoch wesentliche wissenschaftspolitische Aufgaben zu.

Der öffentliche Widerstand gegen gesetzlich erzwingbare Großzählungen der amtlichen Statistik, die Restriktionen des Zugangs zu Individualdaten der amtlichen Statistik für die Sozialwissenschaft und die inhaltliche Verarmung der amtlichen Sozialstatistik haben das Bedürfnis nach Befragungen wachsen lassen, die die Vorteile der amtlichen Statistik mit den Vor-

teilen der Umfrageforschung verbinden und deren Nachteile abmildern (Mayer 1980).

Die ALLBUS-Umfragen stellen eine solche Datenbasis dar. Wie die amtliche Sozialstatistik weisen sie stetige, konventionalisierte Messungen, regelmäßige Wiederholung und Zeitreihenfähigkeit auf. Die Möglichkeit, die Umfragen des ALLBUS zu kumulieren, wird nach wenigen Jahren zu Stichprobengrößen führen, die für sozialwissenschaftliche, nicht regionenbezogene Fragestellungen ähnlich differenzierte Analysen gestatten wie z.B. Mikrozensusdaten. Wie die traditionelle Umfrageforschung erlauben sie eine hohe Flexibilität der Fragengestaltung, wie etwa die Aufnahme von Einstellungsdaten, und eine Datenanalyse auf Individualebene.

Daß sie auf freiwilliger Beteiligung der Befragten beruhen, schafft zwar methodische Probleme in der Behandlung der Ausfälle (vgl. die Beiträge von Kirschner und Hoffmeyer-Zlotnik in diesem Band), stellt aber eine der Lösungen dar, wissenschaftlicher Aufklärung, gesellschaftlicher Selbststeuerungsfähigkeit und dem Schutz der Privatsphäre gleichermaßen Rechnung zu tragen.

Eine weitere wissenschaftspolitische Funktion des ALLBUS-Programms kann nicht hoch genug veranschlagt werden. Die inhaltliche Gestaltung und Auswertung von Bevölkerungsumfragen erfolgt gegenwärtig in der Bundesrepublik überwiegend unter der Kontrolle privater oder staatlicher Instanzen. Dieser einseitigen Verfügungsart kann mit dem "öffentlichen" ALLBUS, wenn auch in kleinem Ausmaß, eine unabhängige, von der Wissenschaft autonom bestimmte Informationsressource entgegengesetzt werden.

Im engeren wissenschaftlichen Bereich zeichnet sich die Möglichkeit ab, daß die Daten des ALLBUS eine gemeinsame empirische Grundlage bilden, auf der Sachkontroversen nicht mehr wegen unterschiedlicher Operationalisierungen und unvergleichbarer, wechselseitig nicht zugänglicher Daten offen bleiben müssen.

Die Entwicklungen seit 1975 haben die Notwendigkeit der Allgemeinen Bevölkerungsumfragen der Sozialwissenschaften aus mehreren Gründen verstärkt.

Nicht nur die Qualität von Umfragen hat sich verbessert, noch sehr viel stärker sind die Kosten von Repräsentativbefragungen gestiegen. Daten-

quellen wie der ALLBUS, die nicht von einzelnen Forschern, sondern von mehr als hundert Benutzern ausgewertet werden, sind daher sehr viel fruchtbarere und ökonomischere Forschungsinvestitionen.

Aber die Umfrageforschung und vor allem die Repräsentativbefragung haben auch aufgehört, der Königsweg der empirischen Sozialforschung zu sein. Andere Analyseeinheiten unter der Bevölkerungsebene, wie z.B. Organisationen, Arbeitsplätze, Schulklassen, Gemeinden, und intensivere, weniger auf Verteilungen gerichtete Erhebungsmethoden sind in den Vordergrund gerückt. Um so wichtiger ist es, repräsentative Umfragedaten in ähnlicher Weise wie die Daten der amtlichen Sozialstatistik als ergänzende, allgemein zugängliche Informationsressourcen zur Verfügung zu stellen.

2. Die Allgemeine Bevölkerungsumfrage als Replikationsstudie

Der methodologische Stellenwert des ALLBUS-Forschungsprogramms oder seines Vorläufers, des inzwischen zehnmal durchgeführten amerikanischen General Social Survey, läßt sich wohl am besten mit Hilfe des Begriffs der "Replikation" und seiner Bedeutungsgeschichte bestimmen.

Der Replikationsbegriff wurde zuerst von dem englischen Biostatistiker Ronald Fisher in die Sprache der Methodologie eingeführt. In seinem Buch "The Design of Experiments" (1935: 8-60, 88, 106) benutzt Fisher diesen Begriff für das Verfahren, eine bestimmte experimentelle Behandlung - z.B. die Anpflanzung einer Getreidesorte - systematisch unter verschiedenen Bedingungen zu wiederholen. So sollte etwa die Aufteilung eines in Quadrate eingeteilten Stück Bodens nach dem Muster eines "Latin Square Design" Variationen im Einfluß der Bodenqualität oder Wetterlage auf den relativen Erfolg ausschließen.

Diese Bedeutung von "Replikation" als systematisch kontrollierte Wiederholung eines Tests zur Bewährung und Bestätigung eines Forschungsergebnisses hat sich noch bis Galtung (1967) erhalten. War Replikation allerdings bei Fisher noch Bestandteil eines strikt experimentellen, varianzanalytischen Designs, so zielt sie bei Galtung eher auf graduelle, kumulative Strategien zur Prüfung der Generalisierbarkeit von Forschungsergebnissen. Randomisierung wird ersetzt durch Heterogenität der Entstehungsbedingungen.

Als eine der Strategien der Sozialindikatorenforschung haben danach Duncan (1969) und Davis (1975) der Replikation eine veränderte Bedeutung gegeben. Mit Hilfe einer "exakten" Replikation von Bevölkerungsumfragen sollte sichergestellt werden, daß gemessene Veränderungen über die Zeit 'wahren' Wandel abbilden. Exakte Replikation soll dabei eine methodische Verfahrensweise heißen, bei der Wiederholungsuntersuchungen folgende Merkmale aufweisen:

1. identische Populationen,
2. identische Stichprobenverfahren,
3. gleiche Frageformulierungen,
4. gleiche Graphik und Technik des Erhebungsinstruments,
5. gleiche Intervieweranleitungen und vergleichbare Interviewer.

Diese Norm der exakten Replikation fordert nicht zuletzt: "In a serious replication, one will repeat faithfully the 'errors' of the original study ..." (Duncan 1969: 27).

Die Norm der Replikation wurde in den ALLBUS-Umfragen praktisch so umgesetzt, daß nur Fragen und Skalen aufgenommen wurden, die bereits in früheren Studien angewandt worden waren und sich insoweit bereits bewährt haben sollten, als dazu nicht nur Daten, sondern publizierte inhaltliche wissenschaftliche Ergebnisse vorhanden sein sollten. Zudem sollten solche älteren Datenquellen für Sekundäranalysen zugänglich sein. Zu diesem Zweck wurde eine nach systematischen inhaltlichen Bereichen geordnete Dokumentation von Instrumenten der empirischen Sozialforschung erstellt, die dann als Auswahlgrundlage für die Fragen des ALLBUS diente.

Es ist nun ohne weiteres einsichtig, daß der Typus der ALLBUS-Umfragen in zweifacher Hinsicht geeignet ist, das Kriterium der Replikation zumindest in dem Sinne von Galtung einzulösen.

Durch Wiederholung zu einem späteren Zeitpunkt (oder beim internationalen Vergleich in einem anderen gesellschaftlichen Kontext) können bekannte Forschungsergebnisse deskriptiver oder ursachenanalytischer Art sich bestätigen und damit in ihrem Verallgemeinerungsgrad verstärkt werden. Die Tatsache, daß solche Bestätigungen an denselben Daten von mehreren Forschern unabhängig vorgenommen werden können, unterstützt diese Zielsetzung von Replikationen noch. Diese Funktion des ALLBUS

wird insbesondere dann, wenn diese Daten für Zwecke der Lehre eingesetzt werden, kaum kontrovers sein.

Maßgeblich für die Anlage des ALLBUS als Replikationsstudie war jedoch die Verfolgung dieser Norm als einer Voraussetzung für eine methodisch unanfechtbare Messung des sozialen Wandels.

Die praktische Einlösung dieser Norm stieß jedoch beim ALLBUS 1980 ebenso wie beim ALLBUS 1982 auf eine Reihe ganz praktischer Schwierigkeiten:¹

a) Eine erste Restriktion hinsichtlich des Replikationsziels stellte das zahlenmäßig geringe Angebot an sowohl methodisch bewährten als auch inhaltlich fruchtbaren Fragen und Instrumenten in der deutschen empirischen Sozialforschung dar, die zusätzlich das Kriterium des freien Zuganges zu den Datensätzen über Archive erfüllt hätten. Eine Ausnahme bildet hier zwar die relativ hohe Kontinuität der Bundestagswahlstudien (1953, 1961, 1965, 1969, 1972, 1976, 1980, 1983), Themen der Wahlforschung brauchten jedoch für den ALLBUS deshalb keine hohe Dringlichkeit einzunehmen, weil die Kontinuität dieser Studien unabhängig vom ALLBUS weiter fortgeführt wird.²

Im Gegensatz zu der Ausgangssituation des General Social Survey in den Vereinigten Staaten sah sich der ALLBUS also in der Münchhausen-ähnlichen Situation, eine auf Zeitreihen zielende Forschungskontinuität begründen zu müssen, ohne sich an gangbare Forschungstraditionen anschließen zu können. Eine erhebliche Schranke bestand auch darin, daß die reichhaltigen Zeitreihen und Originaldaten älterer Umfragen des Allensbacher Instituts für Demoskopie für deutsche Benutzer über Archive nicht frei zugänglich sind.

b) Insgesamt war das Angebot an Einzelfragen sehr viel größer als das Angebot an komplexeren Skalen. Neben einem häufig veralteten Sprachgebrauch scheiterte aber die Replikation von Einzelfragen oft daran, daß deren Indikatorwert unklar blieb. Vor allem die Vorstellung, begrifflich ungeklärte Fragen in Zukunft fortschreiben zu sollen, hat ihre Aufnahme in den ALLBUS verhindert.

c) Von der Replikation ausgeschlossen werden mußten auch solche Instrumente, für welche zum Teil sogar mehrere Erhebungszeitpunkte vorlagen, aber der Problembezug überholt war. Ein Beispiel dazu sind Fragen aus

den fünfziger Jahren zur Geburtenentwicklung, die von dem Problem einer zu rasch wachsenden Bevölkerung ausgehen.

Auch zum Beispiel die traditionsreiche Frage nach der subjektiven Schichteinstufung, die im Fragenprogramm des ALLBUS enthalten ist, gibt Antwortenkategorien vor ('Mittelschicht', 'Arbeiterschicht'), deren Bedeutungsgehalt sich in den vergangenen 30 Jahren vermutlich so sehr verändert hat, daß ein numerischer Zeitvergleich der relativen Anteile in den Selbstzuordnungen erhebliche Interpretationsprobleme aufwerfen dürfte.

d) Ein weitgehend unlösbares Replikationsproblem bieten die technischen Aspekte der geforderten Identität der Stichproben- und Erhebungsverfahren sowie der Identität der Fragebogengraphik und Intervieweranleitungen. In dieser Hinsicht wurde im ALLBUS auf eine strenge Durchsetzung der Replikationsnorm verzichtet, da dies erzwungen hätte, auf den schlechteren methodischen Standard der älteren Umfragepraxis zurückzufallen.

Insgesamt haben die durch die Replikationsnorm gestellten Probleme dazu geführt, daß entgegen der ursprünglichen Absicht mehr Fragen und Instrumente aus Studien der siebziger Jahre in den ALLBUS aufgenommen wurden, die sich als technisch und methodisch besser sowie inhaltlich aktueller und valider erwiesen. Festzuhalten ist auch, daß Fragen zu Verhaltensweisen den strengen Auswahlkriterien besser standhielten als Fragen zu Einstellungen.

Die Replikationsnorm entpuppte sich für den ALLBUS damit vor allem als ein Replikationsdilemma. Im Ergebnis bedeutete dies, daß das ALLBUS-Forschungsprogramm nicht einfach lange Forschungstraditionen regelmäßiger und systematischer fortschreiben konnte, sondern eine solche Forschungstradition selbst begründen mußte.

3. Wege aus dem Replikationsdilemma: Doppelmessungen und 'split-half'-Verfahren

Es gibt jedoch eine Reihe von methodischen Hilfsmitteln, die es gestatten, Methodik und Inhalte von Fragen und Skalen aus früheren Erhebungen beizubehalten, um eine Messung sozialen Wandels überhaupt zu ermöglichen, ohne auf methodisch und inhaltlich gebotene Neugestaltungen ver-

zichten zu müssen. Zu diesen Hilfsmitteln gehören insbesondere die Überprüfung der Auswirkungen von Veränderungen der Fragen- oder Antwort-Vorgaben durch sogenannte 'split-half'-Verfahren und parallele Messungen von Variablen durch ein älteres und ein neues Meßinstrument.

Doppelmessungen sind insbesondere dann sinnvoll, wenn es gute Gründe dafür gibt, eine Zeitreihe nicht mit einem gegebenen Instrument fortzuschreiben, man aber dennoch die Meßreihe nicht völlig neu beginnen will. In den ALLBUS'sen 1980 und 1982 wurde mit dieser Intention die Variable 'subjektiver Status' zum einen mit der von Centers (1949) zuerst verwandten Frage nach der subjektiven Schichtzuordnung gemessen, zum anderen mit einer 'self-anchoring-scale' mit den Skalenwerten 1 bis 10 und den Polbezeichnungen 'Oben - Unten'.

Tabelle 1 zeigt den Zusammenhang zwischen den beiden Meßinstrumenten. Es ist daraus zu ersehen, daß zwar ein relativ enger Zusammenhang der Meßkonstrukte besteht, aber die mit der Variable 'subjektive Schichteinstufung' verbundene Annahme streng disjunkter, hierarchischer Schichten nicht aufrechterhalten werden kann. Insofern bekräftigen die Ergebnisse die Einführung der neuen Skala, da sie den graduellen Charakter sozialer Abstufungen besser erfaßt (vgl. Mayer 1977: 216-232).

Eine zweite Möglichkeit, die Auswirkungen von Veränderungen in Frageformulierungen oder Antwortvorgaben zu überprüfen, liegt in der Anwendung sogenannter 'split-half'-Tests. Dabei wird die Gesamtstichprobe in zwei unabhängige Zufallsstichproben aufgeteilt. Je einer Hälfte der Befragten wird die alte beziehungsweise die veränderte Fragenform vorgelegt. Als Prüfkriterien können dann das Ausmaß der Abweichungen in den Randverteilungen und Abweichungen in den Assoziationen mit Drittvariablen herangezogen werden.

Im ALLBUS 1980 wurden insgesamt drei 'split-half'-Experimente durchgeführt: bei der Frage nach der Wahrnehmung der Verteilungsgerechtigkeit, der Wahrnehmung von Konfliktgruppen und der Struktur von informellen Beziehungen.

Die Motive für die Einführung von Formulierungsveränderungen waren dabei unterschiedlich. Bei der Frage nach der Wahrnehmung der Verteilungsgerechtigkeit hatte der Pretest ergeben, daß der Bezug auf "die angenehmen Dinge des Lebens" Irritationen und Verständnisprobleme auslöste. Die neue Formulierung läßt diesen Bezug weg.

Tabelle 1: 'Oben'/'Unten'-Selbsteinstufung und subjektive Schichteinstufung im ALLBUS 1980 und ALLBUS 1982, in v.H.

'Oben'/'Unten'- Selbsteinstufung		Subjektive Schichteinstufung											
		Insgesamt		Oberschicht		Obere Mittelschicht		Mittelschicht		Arbeiter- schicht		Unterschicht	
		1980	1982	1980	1982	1980	1982	1980	1982	1980	1982	1980	1982
OBEN	10	1.0	0.9	17.6	5.0	1.1	1.2	0.6	0.9	1.1	0.7	0.0	4.5
	9	1.8	1.9	<u>17.6</u>	15.0	6.8	7.8	1.3	1.7	0.6	0.6	3.0	-
	8	8.9	10.8	<u>41.2</u>	<u>65.0</u>	<u>27.4</u>	<u>36.9</u>	8.1	10.5	4.3	3.7	3.0	-
	7	<u>17.2</u>	<u>17.5</u>	11.8	10.0	<u>33.1</u>	<u>32.8</u>	<u>19.9</u>	<u>20.4</u>	7.4	9.5	3.0	-
	6	<u>33.9</u>	<u>33.3</u>	-	5.0	18.3	17.2	<u>41.2</u>	<u>39.7</u>	<u>26.3</u>	<u>28.5</u>	9.1	9.1
	5	<u>19.0</u>	<u>15.0</u>	5.9	-	5.7	2.0	19.0	15.2	<u>24.6</u>	<u>18.5</u>	6.1	13.6
	4	9.6	11.7	-	-	3.8	1.6	7.0	8.6	17.1	19.3	<u>21.2</u>	18.2
	3	5.6	6.3	-	-	3.4	-	2.3	2.5	11.3	13.8	<u>24.2</u>	<u>22.7</u>
	2	1.9	1.9	-	-	0.4	0.4	0.5	0.5	4.5	4.2	18.2	13.6
UNTEN	1	1.0	0.6	3.6	-	-	-	0.1	0.1	2.7	1.2	12.1	18.2
N		2928	2745	17	20	263	244	1645	1511	820	948	33	22

Bei der Frage nach der Stärke von sozialen Spaltungen und Spannungen benutzt die alte Frage den Terminus 'Interessengegensätze'. In den Vor Diskussionen wurde die Vermutung geäußert, 'Gegensätze' seien eine zu schwache Referenz. Die neue Formulierung gebraucht den Terminus 'Interessenkonflikte'.

Bei der Veränderung des Fragenstimulus zu sozialen Netzwerken ging es schließlich darum, einen Wechsel in der wissenschaftlichen Diskussion zu berücksichtigen. Die alte Formulierung nimmt Bezug auf 'Freunde' und damit nur auf emotional positiv besetzte Beziehungen, während die neuere Netzwerkforschung zunächst die Existenz und Intensität von Sozialbeziehungen feststellen will und erst in einem zweiten Schritt nach Art und Zweck von Beziehungen fragt. Zudem lag die Vermutung nahe, daß der 'Freundes'bezug im deutschen Sprachkontext besonders restriktiv aufgefaßt würde.

Die Tabellen 2-5 geben den Vergleich der Randverteilungen wieder.

Tabelle 2: Subjektive Verteilungsgerechtigkeit - Verteilung nach Frageversionen A und B

Antwortkategorie	Frageversion	
	A ("angenehme Dinge")	B
	%	%
gerechter Anteil	66.1	69.1
mehr als gerechter Anteil	4.7	3.0
etwas weniger	24.4	22.8
sehr viel weniger	4.8	5.1
weiß nicht	(65)	(121)
keine Angabe	(4)	(1)
N	(1.510)	(1.445)

Frage 23, Version A:

Im Vergleich dazu, wie andere hier in der Bundesrepublik leben: glauben Sie, daß Sie Ihren gerechten Anteil an den angenehmen Dingen des Lebens erhalten, mehr als Ihren gerechten Anteil, etwas weniger oder sehr viel weniger?

Frage 23, Version B:

Im Vergleich dazu, wie andere hier in der Bundesrepublik leben: glauben Sie, daß Sie Ihren gerechten Anteil erhalten, mehr als Ihren gerechten Anteil, etwas weniger oder sehr viel weniger?

Tabelle 3: Verteilungsmerkmale der zuerst genannten Interaktionspartner, nach Frageversionen

Merkmal	Frageversion	
	A	B ("befreundet")
Durchschnittsalter in Jahren	(40.1)	(41.1)
	%	%
Anteil 1 Person genannt	90.0	86.9
Geschlecht weiblich	51.7	49.1
Nicht-Verwandte	63.4	69.9
Arbeiter	21.9	28.6
erwerbstätig	56.3	58.2
Parteipräferenz: CDU/CSU	43.2	43.5
SPD	47.6	47.7
FDP	4.4	3.0
N	(1.510)	(1.445)

Einleitung der Fragen 9-14:

Version A:

Wir haben jetzt einige Fragen zu den Personen, mit denen Sie häufig privat zusammen sind: Denken Sie bitte einmal an die drei Personen, mit denen Sie am häufigsten privat zusammen sind. Es kann sich dabei sowohl um Verwandte als auch um nicht-verwandte Freunde und Bekannte handeln, nur nicht um Personen, die mit Ihnen im selben Haushalt wohnen.

Version B:

Wir haben jetzt einige Fragen zu Ihrem Bekanntenkreis: Denken Sie bitte an die drei Personen, mit denen Sie näher befreundet sind und mit denen Sie sich am häufigsten treffen. Es kann sich dabei sowohl um Verwandte als auch um Nicht-Verwandte handeln, nur nicht um Personen, die mit Ihnen im selben Haushalt wohnen.

Tabelle 4: Verteilungsmerkmale aller genannten Interaktionspersonen, nach Frageversionen

Merkmal	Frageversion	
	A	B ("befreundet")
durchschnittliche Anzahl der genannten Personen	(2.35)	(2.19)
Durchschnittsalter in Jahren	(39.8)	(40.0)
	%	%
nicht-verwandte Personen	64.1	69.5
Schwiegereltern	1.2	0.8
Erwerbstätige	55.1	57.6
Arbeiter	23.4	26.5
Parteipräferenz: CDU/CSU	42.5	42.8
SPD	46.4	46.2
FDP	4.8	4.0
N	(1.510)	(1.445)

Frageversion siehe Legende zu Tabelle 3

Tabelle 5: Wahrnehmung von Interessenkonflikten (K) und Interessengegensätzen (G) - Antwortanteile nach Frageversionen in v.H.

Konflikt- dimension	Frage- version	Antwortkategorien					keine Angabe
		sehr stark	ziemlich stark	eher schwach	gibt gar keine	weiß nicht	
Links- Rechts	K	38.8	46.0	13.7	1.5	(98)	(1)
	G	39.8	43.1	14.8	2.3	(113)	-
Arbeitgeber- Arbeitnehmer	K	21.1	48.4	27.6	2.8	(55)	(1)
	G	28.3	46.8	22.0	3.0	(63)	-
Volksschüler- Akademiker	K	19.1	35.1	37.3	8.5	(74)	-
	G	28.0	43.9	23.4	4.7	(65)	-
Ohne Kinder- Mit Kindern	K	12.2	24.8	43.9	19.1	(72)	(1)
	G	13.0	28.9	43.6	14.6	(70)	-
Jung- Alt	K	15.0	36.9	39.9	8.3	(35)	(2)
	G	16.6	43.0	33.7	6.8	(38)	-
Arm- Reich	K	30.9	39.1	23.9	6.1	(67)	(2)
	G	39.9	42.2	13.9	4.0	(45)	-
Erwerbstätige- Rentner	K	6.4	16.2	46.1	31.3	(68)	-
	G	8.4	22.9	47.3	21.4	(66)	-
Politiker- Einf. Bürger	K	19.8	32.1	36.5	11.6	(120)	(2)
	G	25.2	40.6	28.8	5.5	(120)	-
Kapital- Arbeit	K	35.4	40.3	20.1	4.3	(94)	(5)
	G	47.4	37.9	11.2	3.5	(89)	-
Gastarbeiter- Deutsche	K	23.1	40.8	31.3	4.8	(68)	-
	G	26.9	43.5	24.5	5.2	(64)	-
Männer- Frauen	K	5.6	21.4	48.6	24.3	(70)	(1)
	G	7.8	21.9	47.2	23.2	(61)	-

Frage 24, Version K:

Es wird oft gesagt, daß es Interessenkonflikte zwischen verschiedenen Gruppen in der Bundesrepublik gibt, z.B. zwischen politischen Gruppen, zwischen Männern und Frauen usw. Die Konflikte sind aber nicht alle gleich stark. Ich will Ihnen nur einige solcher Gruppen nennen. Sagen Sie mir bitte, ob diese Konflikte Ihrer Meinung nach sehr stark, ziemlich stark, eher schwach sind, oder ob es gar keine Konflikte gibt.

Frage 24, Version G:

Es wird oft gesagt, daß es Interessengegensätze zwischen verschiedenen Gruppen in der Bundesrepublik gibt, z.B. zwischen politischen Gruppen, zwischen Männern und Frauen usw. Die Gegensätze sind aber nicht alle gleich stark. Ich will Ihnen nur einige solcher Gruppen nennen. Sagen Sie mir bitte, ob diese Gegensätze Ihrer Meinung nach sehr stark, ziemlich stark, eher schwach sind, oder ob es gar keine Gegensätze gibt.

Die Prozentangaben beziehen sich auf die Gesamtzahl der Befragten ohne 'weiß nicht' und 'keine Angabe'.

Befragte insgesamt: Version K: N = 1.510; Version G: N = 1.445

Im Gegensatz zu verbreiteten Vorstellungen, aber im Einklang mit neueren Untersuchungen (vgl. Schuman/Presser 1977: 1981, Kap.11) sind in den meisten Fällen die Abweichungen gering, das heißt sie bewegen sich sogar im Rahmen des zu erwartenden Stichprobenfehlers.

Dies gilt auch für die Frage nach den Merkmalen von Interaktionspartnern. Allerdings zeigt sich hier, daß der Bezug auf 'Freunde' im Unterschied zu bloßer Interaktionshäufigkeit doch eine andere Zielgruppe eingrenzt. Dementsprechend werden durchschnittlich weniger Freunde genannt: 2.19 vs. 2.35 bei drei möglichen; es werden mehr nicht-verwandte Personen genannt (und weniger Schwiegereltern!).

Eine partielle Ausnahme bildet die Frage nach der Wahrnehmung von sozialen Konfliktgruppen im Vergleich mit der 'split'-Version ('Interessen-gegensätze'). Bei 5 von 11 vorgegebenen möglichen 'Spaltungen' ist ein deutlicher Unterschied zwischen der Wahrnehmung von starken Interessenkonflikten und starken Interessengegensätzen zu beobachten: Arbeitgeber - Arbeitnehmer, Volksschüler - Akademiker, Arm - Reich, Politiker - einfacher Bürger, Kapital - Arbeit. Die beiden Formulierungen scheinen daher in der Tat verschiedene Wahrnehmungen zu induzieren, nämlich zum einen (seltener) offene Konflikte, zum anderen (häufiger) latente Konfliktgruppen.

Die Hilfsmittel der Doppelmessung und des 'split-half'-Verfahrens können dazu dienen, das Dilemma zwischen der Replikationsnorm und veränderten methodischen beziehungsweise inhaltlichen Standards abzumildern. Bei einer begrenzten Befragungszeit sind ihrer Anwendung, zumal bei einer Mehrthemenbefragung, enge Grenzen gesetzt. Für die ALLBUS-Ziele einer gesellschaftlichen Dauerbeobachtung und Analyse des sozialen Wandels wäre es daher fruchtbarer, die Zahl der Meßzeitpunkte zu erhöhen, also etwa von einem zweijährigen auf einen einjährigen Turnus der Umfragen zu wechseln. Dies würde einen größeren Spielraum eröffnen, alte Meßreihen weiterzuführen und gleichzeitig neue beginnen zu lassen.

Freilich setzt dies das Engagement der Sozialwissenschaft für das ALLBUS-Forschungsprogramm und das Interesse der Öffentlichkeit voraus, über den Wandel von Wertorientierungen und Verhaltensweisen nicht nur mit wohlfeilen Aperçus zu spekulieren, sondern sich zur Selbstaufklärung langfristig angelegter Forschung zu versichern.

Anmerkungen

- 1 Vgl. zum Replikationsproblem auch Rolf Porst (1982).
- 2 Vgl. dazu die Dokumentation dieser Studien in der Zeitschrift für Soziologie, Jg.12, Heft 3 und folgende sowie Kaase und Klingemann 1983.

Literatur

- Centers, Richard, 1949: The Psychology of Social Classes. A Study of Class Consciousness. Princeton: Princeton University Press.
- Davis, James A., 1975: The Log Linear Analysis of Survey Replications. In: Kenneth C. Land, Seymour Spilerman (Hrsg.), Social Indicator Models. New York: Russell Sage.
- Duncan, Otis Dudley, 1969: Towards Social Reporting: Next Steps. New York: Russell Sage.
- Fisher, Ronald, 1935: The Design of Experiments. Edinburgh: Oliver and Boyd.
- Galtung, Johan, 1967: Theory and Methods of Social Research. London: Allen & Unwin.
- Kaase, Max und Klingemann, Hans-Dieter (Hrsg.), 1983: Wahlen und politisches System. Opladen: Westdeutscher Verlag.
- Mayer, Karl Ulrich, 1977: Statushierarchie und Heiratsmarkt. In: Johann Handl, Karl Ulrich Mayer und Walter Müller, Klassenlagen und Sozialstruktur. Frankfurt: Campus.
- Porst, Rolf, 1982: ALLBUS 1982. Systematische Variablenübersicht und erste Ansätze zu einer Kritik des Fragenprogramms. Mannheim: ZUMA-Arbeitsbericht Nr.82/12.
- Schuman, Howard und Presser, Stanley, 1977: Question Wording as an Independent Variable in Attitude Surveys. Sociological Methods and Research, 6: 151-170.
- Schuman, Howard und Presser, Stanley, 1981: Questions and Answers in Attitude Surveys - Experiments on Question Form, Wording and Context. New York: Academic Press.
- Zentralarchiv für empirische Sozialforschung der Universität zu Köln und Zentrum für Umfragen, Methoden und Analysen e.V. (ZUMA) (Hrsg.), 1982: Allgemeine Bevölkerungsumfrage der Sozialwissenschaften - ALLBUS 1980. Codebuch mit Methodenbericht und Vergleichsdaten. Mannheim.

DETERMINANTEN DES INTERVIEWER- UND BEFRAGTENVERHALTENS: PROBLEME DER THEORETISCHEN ERKLÄRUNG UND EMPIRISCHEN UNTERSUCHUNG VON INTERVIEWEREFFEKTEN

Hartmut Esser

1. Einleitung

Unter Interviewereffekten versteht man allgemein jede beobachtbare Abweichung von Interviewergebnissen von einem (vermuteten, gemessenen oder erschlossenen) "wahren" Wert des Befragten, insoweit diese Abweichung mit Eigenschaften des Interviewers zusammenhängt. Schon eine nur etwas genauere Betrachtung verweist jedoch darauf, daß sich hierbei mehrere zu unterscheidende Problembereiche überschneiden.

Erstens können Interviewereffekte das Ergebnis bestimmter intentionaler Handlungen des Interviewers selbst sein; seine "interne" Motiv- und Erwartungsstruktur bestimmt - u.a. über die Wahrnehmung von Eigenschaften des Befragten - sein Verhalten (speziell: die Protokollierung einer gewissen "Reaktion" des Befragten); "Fälschungen" und erwartungs- bzw. einstellungsbezogene Vercodungen gehören hierzu. Dann können Interviewereffekte die Folge von Reaktionen des Befragten auf Eigenschaften des Interviewers sein (z.B. als Rollenselektionen in bezug auf gewisse sichtbare Merkmale wie Alter und Geschlecht des Interviewers). Sie können aber schließlich auch das Ergebnis eines - mehr oder weniger langen - Prozesses der wechselseitigen Beeinflussung von Interviewer und Befragtem sein (z.B. die Übertragung von Interviewererwartungen auf den Befragten durch verbale und nonverbale Beeinflussungen; Strukturierung von Interviewererwartungen durch Reaktionen des Befragten beim Interview). Aus diesen sehr unterschiedlichen Möglichkeiten, daß sich ein bestimmter "Effekt" einstellt, wird es zur Erklärung - und zur empirischen Identifikation - notwendig, diese verschiedenen Ursachenkonstellationen systematisch zu unterscheiden.

Nach einer solchen Klärung der verschiedenen Dimensionen von Ursachen für Interviewereffekte wird es dann notwendig, zur Vorhersage und all-

gemeinen Erklärung bestimmter Effekte, und nicht zuletzt zur Integration der vielfältigen und z.T. widersprüchlichen empirischen Ergebnisse zum Problem, einen für Interviewerverhalten, Befragtenverhalten und die "Inter-Aktion" beim Interview gemeinsamen theoretischen Bezug zu finden.

Das Ziel des Beitrags ist es, zunächst die verschiedenen Ursachenkonstellationen für Interviewereffekte genauer zu explizieren, als dies bisher der Fall war (Teil 1). Daran anschließend soll dann eine allgemeine theoretische Grundlage zur Erklärung sowohl des Interviewerverhaltens als auch des Befragtenverhaltens entwickelt werden (Teil 2). Diese allgemeine theoretische Basis soll dann in Teil 3 zur Erklärung der zuvor explizierten unterschiedlichen "Explananda" der Interviewereffekte (Fälschungen, erwartungsgemäße Codierungen, Befragtenbeeinflussungen, Reaktionen des Befragten auf den Interviewer) genutzt werden. Im abschließenden Teil 4 sollen dann die Ergebnisse der allgemeinen theoretischen Erklärung dieser Explananda in bezug auf Möglichkeiten untersucht werden, Kausalmodelle zur Erklärung von Interviewereffekten zu spezifizieren und empirisch zu überprüfen. Ein Ergebnis dieser Analyse ist - es sei vorweggenommen -, daß man ohne die Messung auch der "internen" Persönlichkeitsstruktur der Interviewer aus vorgefundenen empirischen Ergebnissen keinerlei Rückschlüsse über das Vorliegen bzw. die Art von Interviewereffekten ziehen kann. Den Abschluß dieses Teils bildet die Skizze eines "dynamisierten" Modells des Interviewverlaufs, wie es an sich sowohl zur theoretischen Erklärung als auch zur empirischen Untersuchung von Interviewereffekten notwendig wäre; der Untersuchungsansatz im ALLBUS 1980 kommt diesem Standard näher als die überwiegende Mehrzahl auch spezieller Methodenuntersuchungen - die Grenzen der empirischen Überprüfbarkeit werden aus diesem Modell aber auch sehr deutlich werden.

2. Der Interviewprozeß und Dimensionen von Interviewereffekten

Es war eingangs darauf hingewiesen worden, daß eine Analyse von Interviewereffekten die verschiedenen möglichen Verursachungskonstellationen systematisch beachten muß und daß erst von daher eine theoretische Erklärung der empirischen Prozesse möglich ist. Ausgangspunkt der folgenden Explikation der verschiedenen Dimensionen von Interviewereffekten

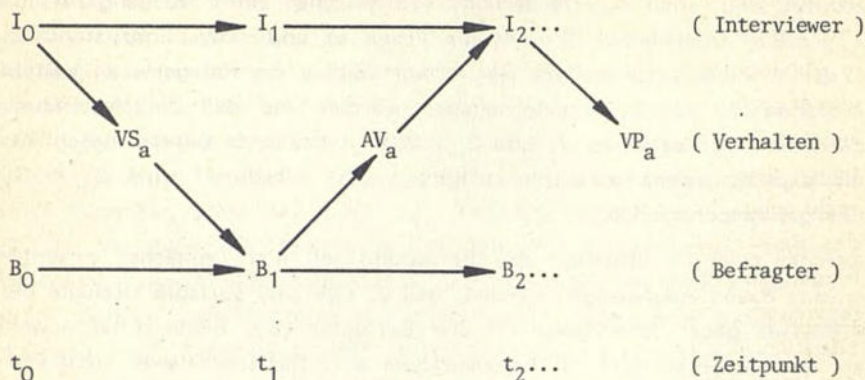
ist eine gemeinhin bekannte Leerformel: Das Interviewprodukt sei das Resultat eines "sozialen Prozesses" und es sei abhängig von Eigenschaften des Befragten wie des Interviewers sowie deren wechselseitiger Beeinflussung (vgl. das Modell bei Kahn und Cannell 1968: 153). Derartige Modelle weisen mindestens zwei Auslassungen auf. Erstens wird in ihnen verdeckt, daß es sich um einen erst noch zu entzerrenden Prozeß einzelner Zustände und Reaktionen handelt, bei dem der Zustand einer Stufe das Explanandum der jeweils vorhergehenden und eine Randbedingung der jeweils folgenden Stufe ist. D.h.: Es wird nicht explizit genug gemacht, daß die Variablen des Modells (demographische Eigenschaften, Motive, Perzeptionen usw.) nicht unmittelbar zum Explanans des Interviewerproduktes gehören, sondern erst im Rahmen einer sogenannten genetischen Erklärung (vgl. Stegmüller 1969: 352ff.) wirksam werden. Zweitens wird nicht explizit gemacht, daß die Interviewer-Befragten-Interaktion im Rahmen auch einer weiteren Situationseinbettung erfolgt und nicht nur von den Eigenschaften der aktuell involvierten Personen abhängig ist. Interviewereffekte wären demnach erst im Rahmen jener anderen Situationsstrukturierungen, von denen der Frageinhalt eines der wichtigsten Elemente ist, einzuordnen. Die Analyse des Interaktionsprozesses im Interview selbst soll im folgenden vorgenommen werden. Die Einordnung von (bestimmten Aspekten von) Interviewereffekten in den Gesamtrahmen des Verhaltens bei Forschungskontakten wird dann im Anschluß an die allgemeine theoretische Grundlegung erfolgen.

Untersuchungen zu Interviewereffekten bestehen in aller Regel - wie in der üblichen substantiellen Forschung - daraus, daß bestimmte Variablen als erklärende Größen für den jeweiligen Effekt vermutet und untersucht werden. Die theoretische Verbindung zwischen den Variablen geschieht dann meist auf der Grundlage von Annahmen über sozialpsychologische Quasi-Gesetze (z.B. Hang zur Selbstpräsentation; Verminderung von Dissonanzen; Wahrnehmung in Stereotypen usw.) und der Vermutung, daß die Variablen jeweils Exemplifizierungen des Prämissenteils der angenommenen Quasi-Theorien seien. Unter gewissen Bedingungen mag dieses Verfahren ausreichen. Es würde z.B. dann ausreichen, wenn mit der vorgeschlagenen Hypothese jeweils auch alle die Zwischenprozesse bekannt wären, die über eine Kette von Bedingungen und Folgen, die dann wieder Bedingungen sind (usw.), den unterstellten bzw. auch beobachteten Va-

variablen-Zusammenhang erklären. Genauer: Zur adäquaten theoretischen Erklärung von Variablen-Zusammenhängen, die das Resultat eines "Prozesses" sind, wird es erforderlich, daß zwischen einem Ausgangszustand Z_0 (z.B.: Interviewer A stellt die Frage a) und einem Endzustand Z_n (z.B. Interviewer A notiert als Antwort auf a die Kategorie k) weitere Zustände $Z_1 \dots Z_{n-1}$ eingeschoben werden und daß die Verbindung zwischen den Zuständen Z_i und Z_{i-1} über n bekannte Gesetze geschieht, die explizit jeweils aufzuführen wären. Erst hierdurch wird Z_0 in Z_n übergeführt und erklärt.

Bezogen auf die Situation der Befragung soll hier (zunächst vereinfachend) davon ausgegangen werden, daß es nur zwei variable Elemente des Prozesses gäbe: Interviewer (I) und Befragter (B). Beide treten sowohl als Reizsender als auch als Reizempfänger auf. Der Interviewer stellt Fragen, erläutert, motiviert, gestikuliert und perzipiert den Befragten und dessen Reaktionen. Der Befragte perzipiert den Interviewer und die von ihm gestellten Fragen und antwortet, erläutert und gestikuliert. Das Handeln des Befragten bestehe - in aller Vereinfachung - zunächst lediglich aus den Reaktionen auf die vom Interviewer ausgehenden Stimuli (AV). Dieses Handeln muß einerseits erklärt werden; Theorien des sozialen Handelns einschließlich der vorgängigen Erklärung von Situationsperzeptionen wären die Erklärungsgrundlage dieses Teils des Vorgangs. Das Handeln des Interviewers bestehe - ebenfalls in aller Vereinfachung - aus zwei Teilen: Dem Aussenden des Stimuli einerseits (VS) und dem Protokollieren der Reaktionen des Befragten andererseits (VP). Beides - so sei postuliert - folgt ebenfalls im Prinzip den Gesetzen des Handelns allgemein. In grober Vereinfachung und unter Vernachlässigung weiterer Zwischenprozesse (z.B. der Stimulusperzeption und kurzfristiger Rückkoppelung) könnte damit der Interviewprozeß wie folgt schematisiert werden.

Schema 1: Modell des Interviewprozesses



Erläuterungen:

$I_0, I_1 \dots$	Interviewerzustand zum Zeitpunkt $t_0, t_1 \dots$
$B_0, B_1 \dots$	Befragtenzustand zum Zeitpunkt $t_0, t_1 \dots$
VS_a	Abgabe des Stimulus zur Frage a
AV_a	Reaktion auf die Frage a
VP_a	Protokollierung (Codierung) der Reaktion auf Frage a
$t_0, t_1 \dots$	Zeitpunkt 1, 2 ...

Es wird angenommen, daß Interviewer und Befragter die Befragung mit einem gewissen internen Zustand beginnen (I_0, B_0). Die Handlungen im Interview beeinflussen nun die internen Zustände in einer Weise, daß der Befragte auf den Fragestimulus "reagiert" und der Interviewer diese Reaktion dann protokolliert. Eine gewisse "Konstanz" der Persönlichkeitsmerkmale auch über die situativen Einflüsse hinweg wird durch die Beziehungspfeile zwischen Interviewer- und Befragtenzuständen zu den jeweiligen Zeitpunkten gekennzeichnet.

Wenn man das Explanandum auf jeder Stufe des Prozesses als alleiniges Resultat der im Modell benannten Randbedingungen auffaßt und die den

Prozeß verbindenden Gesetze in Gesetzen der Wahrnehmung und des Handelns sieht (wobei im Prinzip auch Wahrnehmungen als "Handlungen" gedeutet und erklärt werden können), dann zerfallen die Explananda des Interviewprozesses in drei Einzelaspekte: 1. Die Reaktionen des Befragten in Abhängigkeit von Stimulierungen und Beeinflussungen durch den Interviewer (und die weitere Umgebung; siehe den Residualpfeil) und gewissen eigenen Merkmalen. 2. Die Stimulierungen, Beeinflussungen und Reaktionen des Interviewers in Abhängigkeit von Fragebogen und Befragtenreaktionen, weiterer Situation (wozu vor allem die beobachtbaren Merkmale des Befragten gehören) und gewissen eigenen Merkmalen (z.B. Einstellungen zum Fragegegenstand). Und 3. die Vercodungshandlungen des Interviewers ebenfalls in Abhängigkeit von Befragtenreaktionen, weiterer Situation und eigenen Merkmalen (z.B. einer "Aufgabenorientierung" bei der Vercodung).

Damit wird es möglich, die Interviewereffekte in unterschiedliche Verursachungsdimensionen einzuteilen. Zunächst seien erstens Effekte unterschieden, die sich aus einem "ungültigen" Vercodungshandeln (Aspekt 3) ergeben: "an sich" korrekte und gültige Befragtenreaktionen werden vom Interviewer nicht in eine entsprechende Vercodungshandlung umgesetzt. Hierbei könnte man bewußte Fälschungen von Interviews (oder Einzelfragen) unterscheiden von ungültigen Vercodungen, die sich auf eine mehr oder weniger unbewußte, aber dennoch durchaus intentionale "Organisation" der Befragtenreaktionen (Fehlvercodungen) zurückführen lassen.

Obwohl Fälschungen durchaus nicht unüblich sind, sei ihnen hier nicht weiter nachgegangen, da es sich um ein doch im Prinzip kontrollierbares Problem bei der Interviewerrekutierung und -kontrolle handeln sollte.

Schwerwiegender sind die Fehlvercodungen aus unbewußten Selektivitäten bei der Vercodung einer gegebenen Antwort. In diesem Zusammenhang sind vor allem zwei Aspekte in der empirischen Forschung näher untersucht worden: Erwartungen der Interviewer an den Befragten nach Rollenmerkmalen und die privaten Meinungen der Interviewer zu den einzelnen Fragen. Interviewererwartungen folgen vermuteten Attitüdenstrukturen, Rollenelement-Zusammenhängen und - zuweilen aus einer sozialwissenschaftlichen Ausbildung her bezogenen - Wahrscheinlichkeitszusammenhängen anderer Merkmale bei den Befragten. Hinsichtlich der Auswirkungen von Interviewererwartungen sind die Ergebnisse, die Herbert Hyman (1954: 104ff.) berichtet, immer noch von Bedeutung: Interviewererwartungen, wie sie sich aus dem Befragungskontext ergeben (z.B. bezüglich Geschlechtsrollen oder Statuserwartungen an den Befragten), schlagen sich in der Tat in unterschiedlichem Vercodungsverhalten nieder. Dieses geschieht um so eher und um so stärker, je weniger ein-

deutig die Reaktion des Befragten ist. Solche Erwartungen scheinen - auch neueren Untersuchungen zufolge (vgl. Erbslöh und Wiendieck 1974: 98) - keine allgemeine Gefahrenquelle für Ungültigkeit zu sein, sondern situationell vor allem erst bei einer solchen Ambiguität aufzutreten.

Der Einfluß der Interviewermeinungen zu den einzelnen Fragen wird übereinstimmend als noch unbedeutender eingeschätzt. Hyman (1954: 129f.) spricht ihnen eine allenfalls situationsspezifische und unsystematische Auswirkung zu - erneut vor allem bei Situationsambiguität. Dieses wird auch in neueren Studien bestätigt (vgl. Cosper 1972: 231). Es ist dabei zu beachten, daß wahrscheinlich die Interviewerattitüden mit speziellen Erwartungen an den Befragten zusammenhängen und daß erst über diesen Mechanismus und unter den bei Erwartungen geltenden Bedingungen sich die Interviewereinstellung als systematische Vercodungsselektion bemerkbar macht (vgl. auch Abschnitt 4).

Die theoretische Erklärung für Fehlvercodungen durch Erwartungen (mit dem Spezialfall der Auswirkung von Interviewerattitüden) wird üblicherweise in dissonanztheoretischen Argumenten gesehen: Interviewer organisieren - wie "normale" Akteure - ihre Wahrnehmungen innerhalb eines bereits bestehenden Wahrnehmungsrahmens und es wird dann die Wahrnehmung selektiert (und in eine Vercodungshandlung umgesetzt), die die geringsten Dissonanzen verursacht. Eine evtl. bestehende Orientierung an der Aufgabenerfüllung als Interviewer ("task orientation") wirkt hierbei so, daß eine "Abweichung" von der Aufgabe selbst Dissonanzen verursacht und somit eher unterbleibt - sofern eine bestimmte Wahrnehmung (und entsprechende Vercodung) überhaupt als Abweichung erkannt werden kann: Ambiguität der Reaktion des Befragten ist eine zentrale Bedingung für die Wirksamkeit von Interviewererwartungen.

Zweitens können Interviewereffekte solche Variationen des Interviewergebnisses in Abhängigkeit von Interviewereigenschaften sein, bei denen die Abweichungen vom "wahren" Ergebnis nicht aus einer Falschvercodung, sondern aus einer - mehr oder weniger - bewußten Abweichung des Befragten in Reaktion auf den Interviewer resultiert. Auch hier können zwei Unterfälle unterschieden werden. Einerseits Abweichungen des Befragten, die dieser aufgrund von Beeinflussungen durch das Verhalten des Interviewers vornimmt (Aspekt 2); und andererseits Abweichungen, die sich daraus ergeben, daß der Befragte - ohne jedes eigene Dazutun des Interviewers - auf Merkmale des Interviewers als Bestandteil der Interviewsituation (Aspekt 1) reagiert.

Insgesamt kann man davon ausgehen, daß bewußte Einflußnahmen des Interviewers auf den Befragten eher die Ausnahme als die Regel darstellen (vgl. Hyman 1954: 74). Nichtsdestoweniger gibt es das Problem und zwar eher als Ergebnis non-verbaler Kommunikationen und Verstärkungen bestimmter Befragtenreaktionen. Solche non-verbale und sicher auch nicht bewußt intendierten Beeinflussungen vermutet z.B. Rosenthal (1963) als Grundlage für die von ihm gesondert beachteten Experimentereffekte: Experimentier in sozialpsychologischen Experimenten übertragen ihre Erwartungen (über hypothesengerechtes Vp-Verhalten) auf ihre Versuchspersonen und diese folgen den Experimentern unter bestimmten Bedingungen; nämlich: bei einem guten Verhältnis zwischen Vp und Experimentier in der Situation. Dieses dürfte auch der Grund sein, warum der "Rapport" im Interview keineswegs nur zu eindeutigen Resultaten im Sinne gültiger und zuverlässiger Daten führt. Die Beeinflussungen des Befragten (bzw. der Versuchsperson) sind um so stärker und um so systematischer, je ausgeprägter das Interesse des Experimentiers an der Bestätigung einer bestimmten Hypothese ist. Im Grunde kann man diese Vorgänge auch als Verzerrungen auf der Grundlage von Interviewererwartungen interpretieren. Der Unterschied zu den o.a. Prozessen ist jedoch, daß hier keine Falschvercodung aufgrund von bloßen Erwartungen, sondern eine durchaus korrekte Protokollierung der Befragtenreaktionen vorliegt, zuvor jedoch der Interviewer bestimmte Beeinflussungen (unbewußt und auch durchaus non-verbal) vorgenommen hat, aufgrund deren der Befragte nun den Forschungsstimulus (auch) interpretiert und seine Reaktionen ausrichtet.

Den wohl wichtigsten Aspekt der Interviewereffekte stellen die Auswirkungen von Interviewereigenschaften als Bestandteil der sozialen Situation "Interview" dar. Und dieses deshalb, weil hier - anders als bei Fälschungen, Fehlvercodungen und Befragtenbeeinflussungen - kaum eine Steuerung oder gar Vermeidung möglich ist. Hinweise auf die Wirksamkeit des Interviewers als Situationsbestandteil für den Befragten gibt es sowohl allgemein, als auch speziell nach bestimmten Interviewermerkmalen. Die Ergebnisse sind insgesamt gleicher Art: Die Anwesenheit eines Interviewers führt zur Antwortenanpassung in Richtung auf allgemeine kulturelle Normen; bei speziellen Eigenschaften des Interviewers bzw. bei besonderen Interviewer-Befragten-Konstellationen wird jeweils ein situationspezifisches "looking good" aktualisiert, das sich weniger aus der je spezifischen Situation speist (wie es interaktionistische Theorie vermeint), als es die großen gesellschaftlichen Spaltungen und die hierbei wechselseitig eingeübten Vorbehalte, Schutzreaktionen und Präsentationen wiedergibt.

Die Anpassung an allgemeine kulturelle Normen durch die Interviewerwesenheit ist vor allem bei Vergleichen von persönlichen und schriftlichen Befragungen nachgewiesen worden. Das frühe Ergebnis bei Ellis (1947), wonach etwa beim Interview 37% (angeblich) ihre Mutter in der Kindheit "very dearly" geliebt hätten, dieses aber nur 25% bei einer entsprechenden schriftlichen Befragung angaben, wurde entsprechend auch später immer wieder für andere Fragestellungen in analoger Weise bestätigt (vgl. z.B. Knudsen, Pope und Irish 1967; Wiseman 1972). Der Mechanismus, über den dieser undifferenzierte Anwesenheits-Effekt wirkt, dürfte wohl die bei Interviews im Vergleich zu schriftlichen Befragungen vermutete unmittelbare Konsequenzenbefürchtung vor unerwünschten Antworten sein. Bemerkenswert ist noch, daß derartige Effekte erst dann nachhaltig sind, wenn der Frageinhalt starke Bezüge zur Alltagssituation des Befrag-

ten hat (z.B. Fragen zur Geburtenkontrolle bei Katholiken, Fragen zum Alkoholkonsum bei puritanischen Protestanten).

Spezielle Anwesenheitseffekte sind für eine ganze Reihe von Interviewermerkmalen - z.T. in sehr beeindruckender Form - nachgewiesen worden: Für Rasse - farbige Interviewer erhalten von farbigen Befragten deutlich "radikalere" und weniger geschönte Antworten als weiße Interviewer (Cantril 1944; Bryant, Gardner und Goldman 1966); für Ethnie - orientalische Interviewer erhalten bei Fragen über die Akzeptanz von Orientalen mehr "sozial erwünschte" Antworten als Interviewer anderer ethnischer Zugehörigkeit (Athey, Coleman, Reitman und Tang 1966); für Religion - jüdisch aussehende Interviewer mit jüdischem Namen erhalten systematisch mehr philosemitische Antworten als andere Interviewer (Robinson und Rhode 1946). Ähnliche Effekte sind für Geschlecht (Hyman 1954: 164f.), für Alter (Benney, Riesman und Star 1956) und für soziale Schicht (Katz 1942) nachgewiesen worden.

Die Wirkungsweise dieser Effekte wird durch einige besondere Ergebnisse etwas erhellt. Es kann sich dabei einerseits auch um ungewollte Beeinflussungseffekte durch den Interviewer handeln, bzw. darum, daß erst im Verlauf des Interviews der Befragte gewisse Erwartungen des Interviewers erschließt, und dann darauf reagiert. Es berichtet z.B. Cospers (1972: 231ff.) davon, daß protestantischen Interviewern deutlich weniger Alkoholabhängigkeit zugegeben wurde - und dies besonders bei einer sehr langen Interviewdauer. In der Regel geschieht die Antwortausrichtung aber bereits auf der Grundlage der sichtbaren Kategorisierungsmerkmale der Interviewer, ohne daß diese selbst irgendetwas beeinflussen müßten. Dabei scheinen die Befragten in der Interviewsituation vor allem das Problem lösen zu wollen, mögliche negative Konsequenzen aus bestimmten "wahren" Antworten zu minimieren. Hierfür gibt es drei wichtige Hinweise. Erstens treten die genannten Effekte vor allem und am deutlichsten bei Fragen mit sog. Bedrohungspotential auf (vgl. Williams 1964); hierbei kann, wie auch bei anderen Verzerrungen im Interview, auch ein guter "Rapport" nicht helfen (Williams 1968). Zweitens treten die Effekte dann am deutlichsten auf, wenn es eine klar identifizierte Beziehung des Frageinhalts zu negativen (oder auch positiven) vermuteten Konsequenzen gibt. Dies kann sich daraus ergeben, daß der Interviewer als Hinweis für eine spezielle Öffentlichkeit der Antwort (z.B. einen bestimmten Auftraggeber der Untersuchung) dient. Dies kann sich aber auch auf gesamtgesellschaftliche Spaltungen beziehen, die über die Eigenschaften des Interviewers in die Interviewsituation hineingetragen werden. Schuman und Converse (1968) berichten von einer Untersuchung aus dem Jahre 1942, die in Memphis und in New York durchgeführt wurde. Während in der Südstaaten-Stadt Memphis klare Interviewereffekte nach Rasse erkennbar waren, galt das für New York nicht, wo das Rassenproblem - nach Auffassung der Autoren - weniger dramatisch sich darstellte als in Memphis. Drittens hängen die Effekte von der (vermuteten) "Öffentlichkeit" der Interviewsituation selbst ab: Bei erfolgreichen Anonymitätssicherungen sinken die Effekte auch der speziellen Interviewereigenschaften deutlich (vgl. Hyman 1954: 182ff.); und bei einer unmittelbaren Überprüfbarkeit der "Wahrheit" der Antwort durch den Interviewer sinken die Effekte ebenfalls: "Faktfragen" sind gegenüber "Meinungsfragen" für die genannten Interviewereffekte erheblich weniger anfällig (vgl. Schuman und Converse 1968).

Die theoretische Deutung der zuletzt genannten Effekte als Folge von Befragtenreaktionen auf Interviewereigenschaften knüpft daran an, daß der Befragte den Interviewer als Situationsbestandteil wahrnimmt (wobei diese Wahrnehmung unter Umständen erst durch Beeinflussungen durch den Interviewer selbst zustande kommt). Die Reaktion des Befragten erfolgt dann danach, welche Antwort-"Handlung" innerhalb der Gesamtsituation die relativ geringsten Kosten bzw. den relativ höchsten Nutzen hat. Damit wird für das Befragungsverhalten wie für das Interviewerverhalten ein gleiches Grundmuster unterstellt bei freilich variierenden Bedingungen. Die verschiedenen Effekte wären demnach als Spezialfälle im Rahmen einer allgemeineren Theorie weiter zu explizieren, die soziale Handlungen als situationsorientiertes und rationales Handeln erklären.

3. Eine handlungstheoretische Erklärung der Reaktionen beim Interview

Bevor der Versuch unternommen wird, für die angesprochenen Basishandlungen im Interview (Stimulusabgabe durch den Interviewer, Reaktionen des Befragten, Vercodungshandlungen des Interviewers) in einer speziellen Analyse die genaueren Bedingungen für eine Abweichung vom "idealen" Interviewergebnis anzugeben, muß die allgemeine Grundlage zur Erklärung dieser Handlungen etwas genauer dargestellt werden. Die Handlungen, Reaktionen und Wahrnehmungen der beteiligten Personen im Interview sollen im Rahmen einer allgemeinen Theorie des rationalen und intentionalen Handelns erklärt werden, die - im Prinzip - in der Lage wäre, jedes andere soziale Handeln ebenfalls zu erklären. Aus der Vielfalt der möglichen Basistheorien dieser Art soll dies hier im Rahmen der kognitiven Handlungstheorie erfolgen, wie sie sich aus den Kontroversen zwischen behavioristischen Verhaltenstheorien und mentalistischen Konzepten des teleologischen Handelns in einer Art Synthese herausgebildet haben (vgl. Atkinson 1964; Blalock und Wilken 1979). Danach wählen Personen in einer gegebenen Situation aus der Anzahl der wahrgenommenen Reaktionsmöglichkeiten diejenige aus, von der sie subjektiv annehmen, daß sie am ehesten eine Situation mit der relativ höchsten Motivbefriedigung herbeiführt.

In dieser allgemeinen Konzeption der Handlungserklärung spielen persönliche Faktoren und Merkmale der Handlungssituation eine entscheidende Rolle. Für die personalen Faktoren wird angenommen, daß Personen bestimmte Ziele verfolgen und mit der Realisierung dieser Ziele einen bestimmten "Nutzen" erwarten. Jede Person weist danach einerseits (relativ stabile) Einschätzungen darüber auf, wieviel Nutzen die Verwirklichung eines bestimmten Zieles stiftet. Dieser Aspekt sei hier mit Motivation bezeichnet. Weiter wird angenommen, daß Personen subjektive Erwartungen darüber haben, mit welcher Wahrscheinlichkeit eine bestimmte Handlung in einer gegebenen Situation zur Zielerreichung beiträgt. Dieser Aspekt sei hier mit Erwartung bezeichnet. Die "Tendenz", eine bestimmte Handlung durchzuführen, ergibt sich dann aus dem Produkt von Motivation und Erwartung: $M * E$. Dieses Produkt wird auch als "subjektive Nutzenerwartung" bezeichnet. Man verbindet M und E deshalb multiplikativ, weil man davon ausgeht, daß sowohl bei geringer Motivation wie bei geringer Erwartung (der Zielerreichung durch eine Handlung) die entsprechende Handlung unterbleiben wird. Die gesamte Handlungstendenz (HT) wird allerdings nicht nur von der Motivation und Erwartung in bezug auf ein Ziel bestimmt: Handlungen haben meist auch noch andere - erwünschte oder unerwünschte - Konsequenzen als die Verwirklichung der Ziel-situation. Diese Nebenfolgen werden ebenfalls motivational bewertet; und es wird für sie auch subjektiv geschätzt, wie wahrscheinlich ihr Eintreten bei einer bestimmten Handlung ist. Auch für diese Nebenfolgen gibt es also eine "subjektive Nutzenerwartung". Da Nebenfolgen häufig (keineswegs: immer!) negativen Nutzen haben, wird das entsprechende Produkt aus Motivation und Erwartung für die Nebenfolgen auch als "Kosten" einer Handlung bezeichnet. Wenn "Kosten" vorliegen, bekommt das Produkt für die Nebenfolgen also ein negatives Vorzeichen.

Die gesamte Handlungstendenz für eine Handlung i in bezug auf die Zielkonsequenzen z und die Nebenfolgen k ergibt sich damit als Produktsumme der ziel- und kostenbezogenen Motivationen (M_z, M_k) und der auf die Handlung i bezogenen Erwartungen für Ziel und Kosten (E_{zi}, E_{ki}); bzw. als Summe aus den subjektiven Nutzenerwartungen für das Handlungsziel und für die Nebenfolgen. Formal:

$$HT_i = (M_z * E_{zi}) + (M_k * E_{ki})$$

Bei negativen Nutzenerwartungen für die Nebenfolgen, bei "Handlungskosten", wird HT_i also kleiner, bei positiven Nutzenerwartungen für die Nebenfolgen ("man schlägt zwei Fliegen mit einer Klappe") wird HT_i größer als die Ziel-Nutzenerwartung sein. Man geht nun ganz allgemein davon aus, daß der Akteur aus der Anzahl der ihm bekannten Handlungsalternativen genau die Handlung auswählt, für die HT den (relativ) höchsten (d.h. auch u.U. den am wenigsten negativen) Wert aufweist.

Obwohl man im Grunde zur Bestimmung von HT die Zielkonsequenzen und die Nebenfolgen nicht gesondert aufführen müßte, da HT alleine vom erwarteten "Nettonutzen" abhängt, sei dennoch hier diese Erweiterung vollzogen. Dies hat zwei Gründe: Einmal kann man nun theoretisch erfassen, daß eine geringere Handlungstendenz auf zweierlei Weise zustande kommen kann: Entweder sind bereits die Zielkonsequenzen nur mit geringer Nutzenerwartung versehen (bei Nutzenerwartungen für Nebenfolgen von Null). Dann ist zwar die Handlungstendenz gering, der Akteur handelt aber in "Konsonanz" mit seinen Erwartungen und Einschätzungen. Oder aber, der Nettonutzen ist gering als Folge hoher Ziel-Nutzen-Erwartungen und gleichzeitig fast ähnlich hoher Kosten-Erwartungen. Zwar ist HT hier ähnlich gering wie im obigen Fall; aber der Akteur handelt nun in hoher "Dissonanz" und entsprechendem internen Streß.

Für das Problem der Erklärung von Reaktionen im Interview ist eine solche Unterscheidung von konsonanten und dissonanten Handlungen deshalb von Bedeutung, als im letzteren Fall durch gewisse Situationsänderungen sich die Handlungstendenz drastisch ändern läßt. Z.B. kann man durch Anonymitätszusicherungen die "Kosten" für gewisse "heikle" Antworten vermindern und damit die Handlungstendenz für eine "wahre" Antwort erhöhen. Derartige situationale Eingriffsmöglichkeiten sind für Handlungen mit nur auf ein Ziel bezogenen Nutzenerwartungen ohne Wirkung. Überdies kann der "interne" Streß selbst zu einem "Kostenfaktor" werden, der bei der bloßen Betrachtung des "Nettonutzens" leicht unbeachtet bliebe.

Die Besonderheit dieses Ansatzes ist es, daß er die Erklärung von intentionalem, "reflektiertem" Handeln innerhalb einer nomologischen Konzeption vornimmt, dabei aber gleichzeitig auf den Beitrag eines kalkulierenden "Organismus" Rücksicht nimmt. Dabei wird allerdings auch zunehmend ein Problem der Datenerhebung deutlich: Es sind im Einzelfall sowohl die internen Dispositionen der Akteure (d.h.: ihre "Lerngeschichte") wie die subjektiven Situationselemente zu erheben. Hierin liegt wohl das - berechnigte - Insistieren der "interpretativen" Ansätze begründet, wenn sie die Notwendigkeit einer Einfühlung in die Beweggründe des Handelnden und dessen Situationsdeutungen betonen.

Die Handlungen und Reaktionen im Interview können nach diesem Modell zunächst ganz allgemein interpretiert werden: Ein Interviewer weicht bei seinen Handlungen im Interview dann nicht von seiner "Aufgabe" ab,

wenn er hierzu - z.B. über eine starke "task orientation" - besonders motiviert ist, wenn er über diese Handlung auch die Aufgabe zu erfüllen glaubt und wenn ihm dies andere wichtige Ziele - z.B. den Erhalt eines stabilen Weltbildes, in dem alle alten katholischen Frauen CDU und alle jungen protestantischen Männer SPD wählen - nicht über Gebühr verletzt.

Ein Befragter weicht in seiner Reaktion entsprechend dann nicht von einer "gültigen" Antwort ab, wenn die "gültige" Reaktion in Verbindung zu hoch bewerteten Folgen steht und wenn diese Reaktion keine allzu hohen Kosten (z.B. Mißbilligung durch den Interviewer) verursacht.

Die personalen Nutzenschätzungen und subjektiven Erwartungen reichen jedoch nur in Ausnahmefällen (z.B. bei einer hohen situationsunspezifischen "Wertrationalität") aus, um ein bestimmtes Handeln in einer Situation zu erklären. Es wird zweitens notwendig, die jeweiligen situationalen Gegebenheiten zu berücksichtigen. Personen nehmen danach die Beurteilung von Handlungsalternativen vor dem Hintergrund einer Wahrnehmung der Situation vor (vgl. dazu auch den Beitrag von Schanz und Schmidt 1983 in diesem Band). Hierbei sind jedoch einige Differenzierungen vorzunehmen.

Gegebene Situationen unterscheiden sich erstens darin, ob sie vom Akteur überhaupt als für die Zielkonsequenzen bedeutsam angesehen werden: Nicht in allen Situationen stehen etwa soziale Anerkennung, Karriere oder das ewige Heil auf dem Spiel. Dieser situationale Faktor sei hier mit Relevanz bezeichnet. Weiter unterscheiden sich Situationen danach, inwieweit überhaupt festgestellt werden kann, ob sie von Bedeutung für die Ziele des Akteurs sind bzw. welche Handlungen als zielrelevant gelten können. Wenn fremde Personen in einer undefinierten Situation aufeinander treffen, ist - z.B. bei einer Befragung - nicht klar, auf welche Weise z.B. "soziale Anerkennung" zu erlangen wäre (selbst wenn dieses in der Situation von höchster "Relevanz" wäre). Dieser Aspekt sei hier mit Transparenz bezeichnet. Relevanz und Transparenz bestimmen in multiplikativer Verknüpfung, ob in einer Situation Zielmotivationen und -erwartungen überhaupt handlungsrelevant werden können: Nur wenn sowohl Relevanz und Transparenz einer Situation s in bezug auf eine Handlung i vorliegen, können auf die Handlung bezogene Nutzenerwartungen in der Situation bedeutsam werden. Zur Bestimmung der situationalen "Kraft" für eine Handlung i in der Situation s muß also zu den personalen Nutzenerwartungen ($M_z * E_{zi}$) noch das Produkt aus Relevanz für die Situa-

tion s (R_s) und Transparenz der Situation s (T_s) multiplikativ hinzugefügt werden.

Es muß noch ein weiterer situationaler Aspekt ergänzt werden: Wenn eine Situation so beschaffen ist, daß in ihr nur zielbezogene und keinerlei sonstige (negative) Konsequenzen zu erwarten sind, würden nur die zielbezogenen Nutzenerwartungen (zusammen mit der Wahrnehmung von Relevanz und Transparenz der Situation) handlungsbestimmend sein. Dieser Aspekt, daß Situationen unterschiedlich von (negativen) Nebenfolgen abgeschottet sind (bzw. wahrgenommen werden), sei mit Isolierung der Situation bezeichnet. Der Wert der wahrgenommenen Isolierung einer Situation für Nebenfolgen (Null bei hoher Isolierung; 1 bei nicht vorhandener Isolierung) wird dann zu den Kostenerwartungen ($M_k * E_{ki}$) multiplikativ hinzugefügt: Nimmt die Isolierung der Situation s den Wert 1 an, dann kommen die Kostenerwartungen voll zum Tragen, nimmt I_s einen Wert von Null an, dann werden die Kostenerwartungen nicht handlungsrelevant. Da die Isolierung situationell sehr variieren kann, war oben auch eigens nach Ziel- und Kostenkonsequenzen unterschieden worden.

Da die Kostenaspekte einer Handlung (und damit die Isolierung einer Situation von Nebenfolgen) erst dann überhaupt handlungsbedeutsam werden, wenn die Situation selbst Relevanz und Transparenz aufweist, muß zu den Kostenerwartungen neben der Isolierung selbstverständlich auch noch R_s und T_s multiplikativ hinzugefügt werden. Die gesamte Formel zur Bestimmung der resultierenden "Kraft" K_{is} für die Wahl einer Handlung i in einer Situation s ergibt sich damit wie folgt:

$$K_{is} = (M_z * E_{zi})(R_s * T_s) + (M_k * E_{ki})(R_s * T_s * I_s) .$$

Zu bemerken ist weiter, daß mit den Variablen Relevanz, Transparenz und Isolierung lediglich die Größe "Kognition" der Situation weiter ausdifferenziert wird, die üblicherweise in ähnlichen Ansätzen zur Erklärung von Handlungen herangezogen wird (vgl. den Ansatz von Schanz und Schmidt 1983, in diesem Band). Diese Ausdifferenzierung wird ebenfalls - ähnlich wie die Berücksichtigung des Kostenfaktors - erforderlich, um Variationen in der Situationswahrnehmung durch den Akteur genauer auf Merkmale der Akteure (z.B. unterschiedliche Fähigkeiten zur Situationserkennung) und der Situationen (z.B. variierende "objektive" Grade an Situationsambiguität) beziehen zu können.

Welche Handlung nun der Akteur in einer gegebenen Situation wählt, hängt dann von der Entscheidungsstrategie des Akteurs ab. Üblicherweise

wird angenommen, daß die Handlung mit dem relativ größten K_{is} -Wert gewählt wird. Hier wird - um diese Art der Handlungserklärung dann statistisch in "Pfadkoeffizienten" mit unterschiedlichem "Vorzeichen" und unterschiedlicher Stärke ausdrücken zu können - davon ausgegangen, daß sich die verschiedenen möglichen Handlungen nicht nur in der Kraft, sondern auch in der "Richtung" von einander unterscheiden: Diese Richtung ergibt sich dabei aus den subjektiven Erwartungen, welche genaue Handlung in einer Situation zur Zielerreichung notwendig wäre. Für die unterschiedlichen, in einer Situation als "relevant" wahrgenommenen Ziele und Orientierungen des Akteurs ergeben sich demnach ggf. unterschiedliche Handlungsrichtungen. Z.B. würde sich zur Erfüllung einer "Norm zur Wahrhaftigkeit" als Reaktion auf eine Interviewfrage ein "ja", zur Erreichung "sozialer Anerkennung" aber möglicherweise ein "nein" als geboten erweisen können.

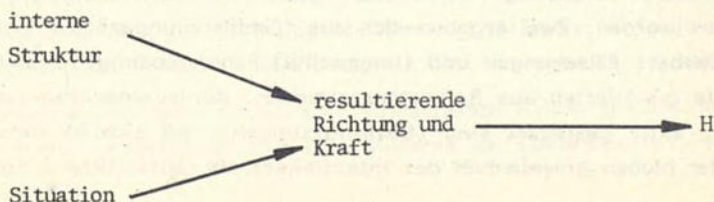
Die Richtungen bestimmter Handlungen (in den subjektiven Erwartungen) können übersituationell relativ stabil sein (z.B. bei gewissen Primärbedürfnissen, oder bei institutionell stark kontrollierten Handlungsabläufen oder auch bei hoher "Wertrationalität"). Sie können sich aber auch - v.a. in "nicht-definierten" Situationen - erst in der Situation selbst als Folge eines "Aushandelns", einer wechselseitigen Anpassung an die "Intentionen des Anderen", einer "Definition der Situation" durch die Akteure ergeben. Eine solche situationale Handlungsausrichtung darüber, welche Handlung in einer Situation welches Ziel bedient, sei hier als "situationales looking good" bezeichnet.

Die tatsächlich resultierende Handlung H in einer Situation ergibt sich dann als ein "Kompromiß" aus allen, in einer Situation nach Richtung und Kraft eingeschätzten Handlungsalternativen (nach Art der Addition von Vektoren mit unterschiedlicher Richtung und Länge). Das heißt: Richtung und Kraft der verschiedenen wahrgenommenen Alternativen kombinieren sich zu einer "resultierenden" Richtung und einer "resultierenden" Kraft in bezug auf eine bestimmte Handlung. Resultierende Richtung und Kraft (für eine bestimmte Handlung, z.B. einstellungsgemäß zu codieren, einen Befragten nicht zu beeinflussen, auf eine Frage mit "ja" zu antworten) "verursachen" somit die Handlung (H).

Wenn man nun davon ausgeht, daß bestimmte Persönlichkeitsmerkmale (Motivationen, Erwartungen, Einstellungen, Präferenzen) bzw. indirekt darüber verbunden bestimmte "Variablen" (wie Alter, Geschlecht, Soziale Schicht) mit einer bestimmten Richtung und Kraft jeweils in bezug auf eine "abhängige" Handlung zusammenhängen, und daß sich aus der Kombination dieser verschiedenen "Variablen" eine resultierende Richtung und Kraft ergibt, dann kann man das "Vorzeichen" der Richtung und die "Höhe" der Kraft über Vorzeichen und Höhe z.B. von Pfadkoeffizienten schätzen. Diese substantiell-theoretische Verbindung und "kausale" Interpretation von Pfadkoeffizienten ist aber erst dann möglich, wenn im Hintergrund eine explizite theoretische Erklärung steht. In ähnlicher Weise kann man dann die "Wirkung" von Situationsvariablen in ein statistisch formuliertes Kausalmodell eingehen lassen.

Wenn man den gesamten Vorgang vereinfacht, dann kann man die resultierende Richtung und Kraft als von zwei Faktoren abhängig ansehen: Die "interne Struktur" des Akteurs (seine Motivationen und Erwartungen sowie seine Fähigkeiten und Neigungen, Situationen wahrzunehmen); und die gegebenen bzw. wahrgenommenen Situationsmerkmale (aus denen sich Relevanz, Transparenz und Isolierung ergeben). In einer noch sehr groben Sicht kann man weiter davon ausgehen, daß die interne Struktur einerseits die Situationsmerkmale beeinflussen kann (z.B. durch unterschiedliche Fähigkeiten zur Situationserkennung) und daß andererseits sich über Situationsmerkmale auch die interne Struktur (z.B. die subjektiven Erwartungen im Verlauf einer Situationsdefinition) verändern kann. Dieses sei hier einstweilen als eine "Korrelation" von interner Struktur und Situation gekennzeichnet. Zur Erklärung einer bestimmten Handlung könnte man damit das folgende Kausalschema konstruieren (Schema 2).

Schema 2: Kausalmodell der Handlungserklärung



Die "Verursachung" der Handlung durch die interne Struktur (z.B. eine bestimmte Motivation) ergäbe sich damit als "indirekter kausaler Effekt" zwischen interner Struktur, resultierender Richtung und Kraft und Handlung, sowie als korrelativer Effekt über die Situation. In analoger Weise ist die Situation mit der Handlung nur indirekt verbunden. Kausal-koeffizienten zwischen interner Struktur bzw. Situation und einer Handlung verkürzen damit im Grunde immer einen tatsächlich indirekten Bezug. Dennoch: Es wird auf diese Weise im Prinzip möglich, eine "handlungstheoretische" Erklärung von Handlungen in Kausaldiagramme zu überführen.

Es sei noch ergänzt, daß nun auch deutlich wird, auf welcher indirekten Weise "in Wirklichkeit" solche Basisvariablen wie Alter, Geschlecht und soziale Schicht mit einer bestimmten Handlungsneigung nur verbunden sind. Alter, Geschlecht und Schicht könnten einerseits als Indikatoren oder andererseits als "Ursachen" für bestimmte interne Strukturen der Akteure gelten. Damit verlängert sich die Kette der indirekten Beziehungen zwischen diesen Basisvariablen und der zu erklärenden Handlungsweise um insgesamt zwei Pfade (Basisvariable auf interne Struktur, interne Struktur auf resultierende Handlungsrichtung und -kraft und von dort auf die Handlung). Die nur geringen Beziehungen zwischen "demographischen" Variablen und bestimmten abhängigen Verhaltensvariablen werden damit gut erklärbar.

Der skizzierte allgemeine Ansatz soll im folgenden in bezug auf das Interviewer- und Befragtenverhalten etwas detaillierter behandelt werden; daran anschließend wird dann die vorgeschlagene Umsetzung der handlungstheoretischen Erklärung in Kausalmodelle zur Erklärung von Interviewereffekten angewandt. Gewisse Vereinfachungen und Verkürzungen werden dabei (auch aus Raumgründen) nicht zu vermeiden sein.

4. Interviewereffekte im Rahmen der allgemeinen Erklärung von Reaktionen im Interview

Ausgehend von den Basishandlungen im Interviewprozeß (Stimulusgabe, Reaktionen, Vercodung) waren vier Typen von Interviewereffekten identifiziert worden: Zwei ergaben sich aus "Fehlhandlungen" der Interviewer unmittelbar; Fälschungen und (ungewollte) Fehlvercodungen. Zwei weitere Effekte resultierten aus Befragtenreaktionen, die im einen Fall vom Interviewer aktiv gesteuert sind (Beeinflussungen) und sich im anderen Fall aus der bloßen Anwesenheit des Interviewers als (zusätzliches) Situations-

merkmal ergeben. Zu erklären werden daher einerseits drei Aspekte des Interviewerverhaltens sein (Fälschungen, Fehlvercodungen, Befragtenbeeinflussungen) und andererseits der Beitrag des Interviewers als Situationselement für das Befragtenverhalten.

Dazu sei im Anschluß an das o.a. allgemeine Modell der Handlungserklärung ein Modell des Interviewerverhaltens zur Analyse von Fälschungen, Fehlvercodungen und Beeinflussungen sowie ein Modell des Befragtenverhaltens entwickelt. Bei letzterem wird - anders als bei der Erklärung des Interviewerverhaltens - ausführlich auf den Satz der anderen Komponenten der Befragtenreaktion - Fragestimulus, kulturelle Eingebundenheit, Persönlichkeitsmerkmale - einzugehen sein.

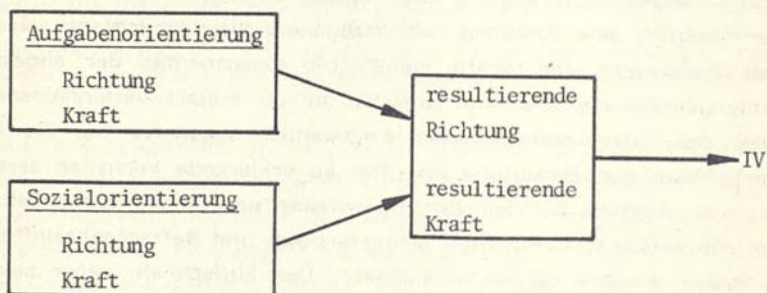
4.1 Interviewerverhalten

Die o.a. allgemeine Erklärung des Handelns postuliert, daß ein Akteur in einer Situation eine Handlung nach Maßgabe eines Kompromisses verschiedener Richtungen und Kräfte wählt. Die Komponenten der allgemeinen Handlungserklärung seien nun zunächst auf ein - stark vereinfachendes - Modell des Interviewerverhaltens angewandt, soweit es für die Interviewereffekte von Bedeutung ist. Das zu erklärende Verhalten seien die drei o.a. Aspekte der auf den Interviewer unmittelbar zurückgehenden Interviewereffekte: Fälschung, Fehlvercodung und Befragtenbeeinflussung (im Modell allesamt mit IV bezeichnet). Der Einfachheit halber seien lediglich zwei unterschiedliche personale Orientierungen in bezug auf IV angenommen (für die selbst im Modell keine Erklärung vorgenommen wird). Diese beiden Komponenten seien in Anlehnung an die Unterscheidung von Hyman (1954: 74) mit Aufgabenorientierung (task involvement) und Sozialorientierung (social involvement) beschrieben. Damit wäre die grobe Richtung der beiden in der Situation vor allem bedeutsamen Handlungstendenzen umrissen: Aufgabenorientierung richtet das Handeln auf ein korrektes Verhalten (keine Fälschungen, keine Fehlvercodungen, keine sachfremden Beeinflussungen). Der motivationale Hintergrund mag die Internalisierung einer Professionsethik des Interviewers o.ä. sein.

Sozialorientierung kann davon eine abweichende Richtung bedeuten. Der motivationale Hintergrund mag ein Interesse am Geldverdienen, an Kon-

takten mit Personen allgemein, ein Interesse an der Bestätigung bestimmter Hypothesen o.ä. sein. Welches Verhalten gewählt wird, hängt dann - unter anderem - von der genauen Richtung der Sozialorientierung ab; denn diese kann u.U. in der Richtung der Handlungswahlen mit der Aufgabenorientierung zusammenfallen (wenn z.B. nicht die Hypothesenbestätigung, sondern das Einhalten methodischer Standards auch das "private" Interesse eines Experimenters ausmacht). Bei Abweichung der Orientierungsrichtungen werden dann weiter die jeweiligen "Kräfte" bedeutsam; genauer: die jeweiligen Nutzenerwartungen sowie Relevanz, Transparenz und Isolierung der Situation in bezug auf entsprechende - aufgabenorientierte bzw. sozialorientierte - Handlungen. Dies ergäbe das folgende - stark vereinfachende - Modell (Schema 3).

Schema 3: Modell des Interviewerverhaltens



Das Modell erfordert einige Erläuterungen. Relativ unproblematisch ist die inhaltliche Bestimmung der Komponente Aufgabenorientierung. Die Richtung ist durch die Aufgabenbeschreibung vorgegeben; mindestens: Vermeiden von Fälschungen, Fehlvercodungen und Befragtenbeeinflussungen. Die Kraft für aufgabenorientiertes Handeln ist jedoch von einer Vielzahl einzelner situationeller Gegebenheiten abhängig, von denen hier einige exemplarisch aufgezeigt seien. Die Relevanz des aufgabenorientierten Handelns dürfte bei Forschungsinterviews fast ausschließlich von "Internalisierungen" der Interviewerrolle abhängen. Erst bei Interviews, die auch

andere Zwecke als die Datensammlung verfolgen, dürfte es hier situationelle Variationen geben. Hinsichtlich der Transparenz dürfte es bei Forschungsinterviews Unterschiede für die Aufgabenerfüllung geben. Dieser Aspekt war bei der Erklärung von Erwartungseffekten mit der Variablen "Ambiguität" der Situation ja auch nachhaltig nachgewiesen worden. Die Isolierung der Situation vor allem von Verletzungen der Aufgabenerfüllung dürfte u.a. von den Kontrollen der Interviewerarbeit abhängen. D.h.: die Aufgabenorientierung ist um so stärker, je besser der Interviewer seine Aufgabe internalisiert hat und je relevanter eine bestimmte Aufgabenerfüllung für sein Rollenverhalten (z.B. bei reinen Forschungsinterviews) ist, je eindeutiger die Handlungsanweisungen definiert sind und je kontrollierbarer seine Aufgabenerfüllung insgesamt ist.

Die Erläuterung der Sozialorientierung soll jeweils gesondert für die drei Abweichungen: Fälschung, Fehlvercodung und Befragtenbeeinflussung erfolgen. Die Richtung der Sozialorientierung für das Interviewerverhalten ergibt sich aus den "aufgabenfremden" Zielen des Interviewers (z.B. rasches Geldverdienen, Vermeidung von kognitiven Dissonanzen aus verletzten Weltbild-Erwartungen, Bestätigung von Hypothesen in Experimenten) und den Vorstellungen über Handlungsweisen zur Verwirklichung dieser Ziele (z.B. Fälschen von Interviews, erwartungsbezogene Fehlvercodungen, Übertragung von Hypothesenerwartungen auf Versuchspersonen). Die Kraft für sozialorientiertes Interviewerverhalten bestimmt sich ebenfalls inhaltlich sehr unterschiedlich für jede der genannten Abweichungen. Die Höhe des Interviewentgeldes (in Kombination mit der jeweiligen Geldnot), die Strukturiertheit der Anhaltspunkte für "Erwartungen" an den Befragten (in Verbindung mit einem bestimmten Weltbild), die Wichtigkeit einer Hypothesenbestätigung in einem speziellen Experiment (in Verbindung mit einem speziellen Interesse an einer bestimmten "wahren" Theorie) mögen jeweils die Relevanz der Sozialorientierung bestimmen. Die Transparenz ergibt sich z.B. entsprechend aus der situationell erkennbaren Wahrscheinlichkeit, daß gefälschte Interviews auch belohnt werden (und nicht z.B. die nachgewiesene Interviewzeit), daß durch eine Vercodungshandlung das Weltbild der Erwartungen auch erhalten bleiben kann, daß Versuchspersonen überhaupt in bestimmter Weise sich beeinflussen lassen. Schließlich werden entsprechende abweichende Handlungen erst dann auch eine stärkere Tendenz erhalten, wenn Isolierung gegeben

ist, d.h. wenn evtl. negative Konsequenzen ausgeschlossen werden können: Entlarvung von Fälschungen durch Kontrollen, Kontrollinterviews zur Aufdeckung von offenbaren Fehlvercodungen, Replikation von Experimenten z.B. zur Aufdeckung von Experimentereffekten.

Es wäre im Prinzip möglich, für die verschiedenen Bestandteile der beiden Orientierungsvektoren ihrerseits erklärende - personale und situationale - Variablen - z.B. Art der Interviewerschulung, Merkmale der Befragten, Organisation und Zusammensetzung des Interviewerstabes, Persönlichkeitsmerkmale der Interviewer - anzugeben. Dieses soll hier (einstweilen) unterbleiben.

Es dürfte jedoch klar geworden sein, daß und wie im Prinzip das Interviewerverhalten als Basisbestandteil des Interviewerprozesses und als Hintergrund eines Teils der beobachtbaren Interviewereffekte auf der Grundlage einer allgemeinen Theorie - unter Hinzuziehung spezieller und jeweils oft nur situationsspezifisch angebbarer "Randbedingungen" - erklärt werden kann. Eine Einordnung und Integration der verschiedenen Ergebnisse der empirischen Forschung könnte von hier ihren Ausgang nehmen.

4.2 Befragtenverhalten

Interviewereffekte treten nicht nur als Folge des Fehlverhaltens von Interviewern auf, sondern auch als Folge der Anwesenheit des Interviewers und der hier durch den Befragten nahegelegten Situationsdefinition. Eine Abschätzung der Bedeutsamkeit derartiger Effekte kann - ähnlich wie beim Interviewerverhalten - nur im Zusammenhang einer Erklärung des Befragtenverhaltens insgesamt, d.h. auch im Zusammenhang der jeweils anderen Determinanten des Befragtenverhaltens erfolgen.

Das Modell zur Erklärung des Befragtenverhaltens (Schema 4; vgl. auch Esser 1982) knüpft an ein konkretes Beispiel zur Erklärung einer inhaltlich bestimmten verbalen Reaktion an (hier: Vorurteile gegenüber Minderheiten). Ein bestimmtes Antwortverhalten (AV) kann - ähnlich wie das Interviewerverhalten - nicht unabhängig von inhaltlichen Orientierungen untersucht werden. Dazu wird - in einer weiteren Vereinfachung - davon ausgegangen, daß bei der Stimulusaufnahme und -verarbeitung keine wei-

teren - z.B. kognitiven - Probleme auftreten und daß der Befragte eine bestimmte "wahre" Einstellung besitzt, die bei Fehlen anderer Faktoren das AV vor allem bestimmen würde (z.B. eine Einstellung zu Minderheiten, wonach eine entsprechende Frage nach "Vorurteilen" mit "ja" oder "nein" beantwortet würde). Diese "anderen" Faktoren seien hier drei Komplexe, die jeweils eigene Handlungsorientierungen ("Richtungen") für das AV hervorbringen.

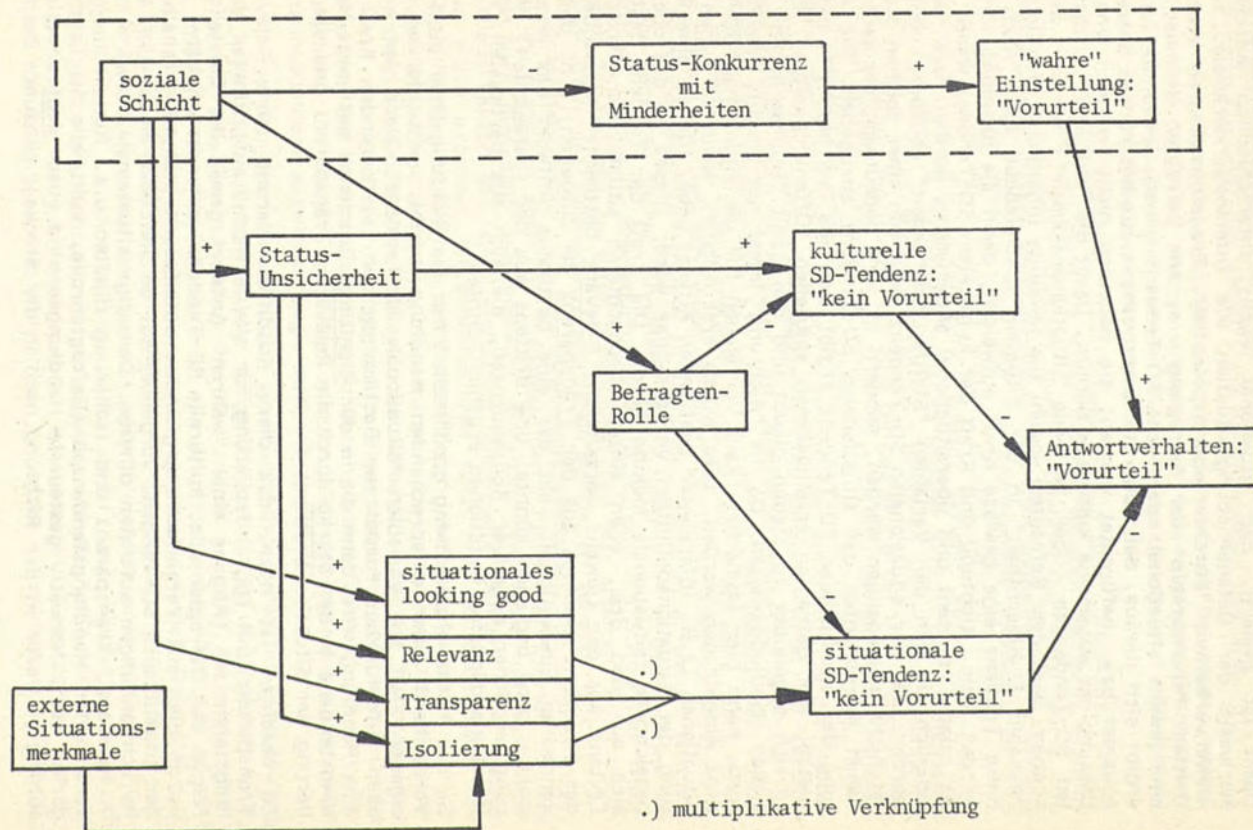
Diese drei Komplexe bestehen einmal aus einer Tendenz, das AV an allgemeinen (internalisierten) kulturellen Normen auszurichten, wie sie entweder gesamtgesellschaftlich oder auch subkulturell verschieden für bestimmte Themenbereiche ausgebildet sind. Dieser Komplex sei hier kulturelle SD-Tendenz (von "social desirability") genannt. Zweitens sei das AV von situationsspezifischen Erwartungen und Befürchtungen bestimmt, wie sie z.B. durch den Interviewer nahegelegt und aktualisiert werden. Dieser Komplex sei hier situationale SD-Tendenz genannt. Hierunter werden die o.a. situationsspezifischen Interviewereinflüsse gefaßt. Und schließlich werde drittens das AV auch von einer besonderen Aufgabenorientierung des Befragten im Interview, analog zu der des Interviewers, bestimmt. Dieser Komplex sei hier Befragtenrolle genannt. Das Antwortverhalten wird dann - so lautet die allgemeine Orientierungshypothese - als Resultat des kombinierten Wirkens der vier Orientierungsrichtungen: wahre Einstellung, kulturelle SD-Tendenz, situationale SD-Tendenz und Befragtenrolle erklärt. Genauer: Für jeden der vier Komplexe gibt es eine je spezifische Richtung und Kraft der Reaktionstendenz, deren irgendwie geartete Kombination in die beobachtbare Antworthandlung mündet. Der - hier besonders interessierende - spezifische situationale Interviewereffekt wäre damit als Teil der situationalen SD-Tendenz lokalisiert; seine Auswirkungen sind von dieser, aber auch von der Richtung und Stärke der anderen Antwortdeterminanten abhängig.

Der erste Komplex, die "wahre" Einstellung, wird im Modell (Schema 4) im Rahmen einer (einfach gehaltenen) substantiellen Kerntheorie erklärt. Diese Kerntheorie geht davon aus, daß diskriminierende Einstellungen kausal von einer Status-Konkurrenz mit Minderheiten abhängen, da der Verweis auf askriptiv begründete Überlegenheiten eine Sicherung des bedrohten Status bedeuten könnte. Die Status-Konkurrenz sei dann weiter von der Schichtzugehörigkeit abhängig. Und zwar gebe es eine negative Beziehung zwischen sozialer Schicht und Status-Konkurrenz, da nur die unteren Schichten mit ethnischen Minderheiten um Arbeitsplätze u.a. konkurrieren. Weiter werde keine kausale Beziehung innerhalb der Kerntheo-

rie angenommen. Für das AV wird nun angenommen, daß es in einer nicht-deterministischen Weise von der im Kontext der Kerntheorie erklärten Variablen "wahren Einstellung" abhängt. Es sei hier darauf verzichtet, die Vermittlung zwischen Einstellungen und verbalem Verhalten auch im Rahmen der o.a. Grundtheorie und über die Variablen der Richtung und Kraft bzw. Relevanz, Transparenz und Isolation weiter zu präzisieren. Es wird lediglich davon ausgegangen, daß diese Beziehung nicht-deterministisch ist, wobei der entsprechende Dependenzkoeffizient die jeweilige "Zuverlässigkeit" der Messung ausdrückt. Diese sei hier demnach ausschließlich von Faktoren abhängig, die zwischen "wahrer" Einstellung und AV intervenieren: mangelndes Frageverständnis, Verbalisierungs- und Verständigungsprobleme, Unaufmerksamkeiten, unsystematische Variationen der Erhebungssituation, aber auch: andere Einstellungen, die nicht erfaßt wurden, aber - unerkannt - das AV auch bestimmen. Die Erweiterung dieses Modells der Berücksichtigung unsystematischer Meßfehler zum vollen Modell geschieht dagegen über Annahmen darüber, daß neben der betreffenden abhängigen Variablen der Kerntheorie ("wahre Einstellung") noch andere Faktoren am Zustandekommen des AV beteiligt sind.

Auch bei den kulturellen SD-Tendenzen wird nicht weiter nach den Einzelbestandteilen: Richtung und Kraft, differenziert. Die Richtung liege mit der jeweils verbreiteten (bzw. subkulturell jeweils unterschiedlichen) Norm fest und auch die Stärke solcher "Wertorientierungen" sei relativ invariant, da es ja gerade das wertorientierte Handeln auszeichnet, daß es gegenüber situationellen Variationen sehr stabil ist. Im vorliegenden Beispiel werde angenommen, daß diese kulturelle SD-Tendenz ("social-desirability-Tendenz") einer Äußerung von Vorurteilen entgegensteht; daß also bei Vorliegen dieser kulturellen SD-Tendenz eine etwa bestehende Neigung zur Äußerung von Vorurteilen (auf der Grundlage einer entsprechenden "individuellen" Einstellung) abgeschwächt werde. Es wird nun aber - angeleitet durch entsprechende empirische Ergebnisse zu "social-desirability"-Effekten - angenommen, daß SD-Tendenzen nur bei Vorliegen von Status-Unsicherheit möglich sind; d.h. bei undefinierter Status-Zugehörigkeit (nicht zu verwechseln mit der o.a. Status-Konkurrenz!). Es wird also davon ausgegangen, daß die Stärke der kulturellen SD-Tendenz übersituational stabil von einer durch Status-Unsicherheit handlungsrelevanten (sub)-kulturellen Norm abhängt. Status-Unsicherheit komme aber andererseits erst in mittleren Schichten vor, so daß man eine positive Beziehung sowohl zwischen sozialer Schicht und Status-Unsicherheit, wie zwischen Status-Unsicherheit und kulturellen SD-Tendenzen postulieren kann. Hierüber würden im übrigen die stets beobachteten Zusammenhänge zwischen sozialer Schicht und allgemeinen SD-Tendenzen erklärbar. Diese Beziehungen hätten nun Auswirkungen derart, daß mit steigender sozialer Schicht über Status-Unsicherheit eine SD-Tendenz auftritt, keine Vorurteile zu äußern (selbst wenn sie qua Einstellung vorhanden wären). Im vorliegenden Falle würde eine entsprechende "wahre" negative Beziehung zwischen sozialer Schicht und Äußerung von Vorurteilen zusätzlich verstärkt. Und genau dieses ist gemeint, wenn davon gesprochen wird, daß SD-Tendenzen unter bestimmten Bedingungen (nicht: unter allen Bedingungen) etwaige "wahre" Beziehungen in den auswertbaren Daten inflationieren. Selbstverständlich ist es vor dem Hintergrund einer anderen Kerntheorie auch denkbar, daß eine "wahre" Beziehung invers zu den jeweiligen kulturellen SD-Tendenzen steht. Dann würden die beobachtbaren Beziehungen abgeschwächt werden.

Schema 4: Modell zur Erklärung des Befragtenverhaltens; Äußerung von "Vorurteilen"



Der dritte Komplex behandelt den Einfluß situationaler SD-Tendenzen. Damit ist gemeint, daß situational jeweils unterschiedlich auftretende Merkmale der Datenerhebungssituation wie Interviewermerkmale, Interviewerverhalten, "Furchterregungspotential", Fragethematik, Anwesenheit Dritter, "sponsorship" der Befragung u.a. dem Befragten Hinweise über eine jeweils situational spezifische SD-Tendenz geben. Diese SD-Tendenz ergibt sich daraus, daß konkrete Konsequenzen aus bestimmten Antworten erwartet bzw. befürchtet werden; sie bedeuten nicht lediglich eine Anpassung an allgemeine kulturelle Normen. Damit die "aktuellen" situationalen SD-Tendenzen (über "externe Situationsmerkmale") wirksam werden können, muß dem Befragten sowohl die jeweilige inhaltliche Richtung der "sozialen Erwünschtheit" in der Situation erschließbar sein. Auch muß diese Tendenz eine gewisse Kraft aufweisen, damit sie handlungswirksam werden kann. Richtung und Kraft der kulturellen SD-Tendenz waren oben als relativ konstant und übersituational angenommen worden, bzw. waren lediglich als von der Variablen "Status-Unsicherheit" abhängig vermutet worden. Bei der situationalen SD-Tendenz kann von einer solchen Stabilität nicht ausgegangen werden, sondern es wird bedeutsam, die verschiedenen Komponenten der situationalen SD-Tendenz anzugeben. Die Richtung der situationalen SD-Tendenz ergibt sich aus den jeweils in der Situation vom Befragten erschlossenen speziellen "Öffentlichkeiten", denen er sich gegenüber zu stehen glaubt und den damit "erforderlichen" speziellen Handlungsweisen. Da die "Richtung" immer nur situativ "definiert" wird, kann hier (zunächst) kein Vorzeichen für die situationale SD-Tendenz eingetragen werden. Dies würde erst möglich, wenn es Typen von Situationen (z.B. Interviewer bestimmten Alters) gäbe, für die eine Richtung der Situationsdefinition vorhersagbar wäre. Dieses ist das oben so bezeichnete "situationale looking good". Die Kraft der Tendenz bestimmt sich aus den drei, oben einzeln beobachteten situationalen Faktoren: Erstens aus der (jeweils vermuteten) Relevanz bestimmter Konsequenzen der Situation; zweitens aus der Transparenz der Situation, d.h. der erkennbaren Eindeutigkeit, mit der eine bestimmte Antwort das jeweilige looking good bedienen könnte; und drittens aus der (vermuteten) Isolierung möglicher negativer Konsequenzen, die sich aus bestimmten, nicht situationsadäquaten Reaktionen ergeben könnten.

Da das "situationale looking good" selbst nur die Richtung (hier also: das Vorzeichen) einer entsprechenden Handlung angibt, geht hiervon kein eigener Pfeil zur Variablen "situationale SD-Tendenz". Nach der weiter oben aufgeführten Formel zur Bestimmung der resultierenden Kraft auf eine Handlung wird dann die in der Richtung situationell bestimmte Handlungstendenz in der Stärke durch die Relevanz, Transparenz und die Isolierung der Situation bestimmt.

Zu beachten ist noch, daß diese Ausdifferenzierung nach Relevanz, Transparenz und (ggf.) Isolierung für alle im Modell aufgeführten Handlungstendenzen (Abgabe einer "wahren" Antwort gemäß der Einstellung, Folgen der Befragtenrolle, kulturelle SD-Tendenz) erforderlich wäre. Dies wurde hier aus Vereinfachungsgründen unterlassen und nur deshalb bei der situationalen SD-Tendenz vorgenommen, da hier noch am ehesten starke Schwankungen auftreten dürften. Derartige situationale Schwankungen in Relevanz, Transparenz und Isolierung dürften u.a. für die normativ verankerten Handlungstendenzen (Befragtenrolle, kulturelle SD-Tendenz) gering sein. Normativ gesteuerte Handlungen sind eben nicht situationsabhängig (weder in der Richtung noch in der Stärke); genauer: Bei nor-

mativ orientierten Handlungen können Relevanz, Transparenz und Isolierung (im o.a. Sinne) mit dem Grenzfall 1 angenommen werden, so daß nur der "interne Zustand" des Akteurs die Handlungswahl bestimmt.

Die oben besprochenen Beispiele der Zusammenstellung bei Hyman weisen sämtlich sehr deutliche Muster situationaler Relevanz und Transparenz zusammen mit eindeutig definiertem situationalem looking good und fehlender Isolierung von negativen Konsequenzen auf.

Im Modell werden allerdings situationales looking good, Relevanz, Transparenz und Isolierung nicht als ausschließlich "extern" bestimmte Größen, sondern ihrerseits als (teilweise) durch Variablen der Kerntheorie und der bereits formulierten Methodentheorie bestimmt angesehen. Von sozialer Schicht seien das looking good und die Transparenz abhängig. Es wird angenommen, daß mit steigender sozialer Schicht auch in variierenden Interviewsituationen die subkulturelle Tendenz besteht, keine Vorurteile zu äußern. Die positive Beziehung zwischen sozialer Schicht und Transparenz wird mit der - situationsübergreifenden - besseren kognitiven Ausstattung oberer Schichten begründet, die eine De-Codierung der Situations-Signale erleichtern dürfte. Von der Variablen "Status-Unsicherheit" seien dagegen die wahrgenommene Relevanz situationsadäquaten Verhaltens und die Wahrnehmung der Nicht-Isolierung von Konsequenzen abhängig. Da die Status-Unsicherheit ihrerseits von sozialer Schicht abhängt, wirkt die Variable soziale Schicht somit in direkter bzw. indirekter Weise auf alle vier Bestandteile der situationalen SD-Tendenz. Gemäß den inhaltlichen Postulaten im Modell (die selbstverständlich revidierbar sind) wäre damit (hier) über die situationalen SD-Tendenzen einer der kulturellen SD-Tendenz in der Richtung ähnliche Wirkung zu erwarten: Bestehende "wahre" Neigungen von Personen oberer Schichten, "keine Vorurteile" zu äußern, werden situational verstärkt. Da jedoch die Richtung und die Stärke der situationalen SD-Tendenz zu einem erheblichen Teil von anderen, nicht durch Eigenschaften der Befragten erklärbaren Merkmalen abhängt (die "externen Situationsmerkmale" wie z.B. Interviewerverhalten, sponsorship-bias), kann es durchaus auch zu "gegenläufigen" Tendenzen kommen. Dieses hätte aber nur dann systematische Auswirkungen, wenn ein bestimmter Situations-Faktor bei allen Befragten gleichmäßig auftritt. In welcher Weise diese externen Situationsfaktoren dann auf die situationale SD-Tendenz wirken, kann nur bei Kenntnis der jeweiligen Merkmale als Hypothese eingeführt werden; daher wurde auch kein Vorzeichen eingetragen. Wenn es eine Streuung sehr unterschiedlicher solcher Merkmale über Befragte mit jeweils unterschiedlichen Eigenschaften gibt, dann dürfte sich dieses für evtl. zu schätzende Dependenzkoeffizienten als eine Abschwächung der durch "Richtung" und "Kraft" bestimmten resultierenden Situationstendenz auswirken (vgl. auch Abschnitt 4).

Als vierter Komplex sei schließlich ein häufig unterschätzter (und bislang auch kaum untersuchter) Faktor aufgeführt, von dem angenommen wird, daß er auf eine jeweils vorliegende SD-Tendenz (welcher Art auch immer) abschwächend einwirkt. Gemeint ist die Anbindung des Befragten an eine spezielle Befragtenrolle. D.h. es wird angenommen, daß Befragte auch das methodisch erforderliche Verhalten (verbale Äußerung nach der "wahren" Einstellung) u.U. als "erwünschtes" Verhalten erkannt bzw. internalisiert haben. Auch hier wird nicht weiter nach Richtung und Kraft differenziert: Die Richtung ist vorgegeben und die Kraft variiert - so sei die Annahme - nur nach Maßgabe der Relevanz und Transparenz der Situation für rollengerechtes Verhalten. Da beides - Relevanz und Transpa-

renz - im Beispiel am ehesten bei Mittelschichten vorkommen dürfte, wurden hier positive Beziehungen angenommen. Wie selten eine explizite Orientierung an der Befragtenrolle auch vorkommen mag: Sie ist nicht als eigener Bestandteil der das AV bestimmenden Faktoren auszuschließen. Im Modell wird die SD-korrigierende Wirkung der Befragtenrolle durch jeweils negative Pfade berücksichtigt. Außerdem wird angenommen, daß die Befragtenrolle - als "interessiert-detachiertes Verhältnis" zum Fragegegenstand, zum Interviewer, zur Sozialwissenschaft, zur umgebenden Gesellschaft - am ehesten in mittleren und oberen (dagegen nicht: in obersten!) Schichten verinnerlicht ist bzw. erfüllt werden kann. Das heißt, daß es eine positive Beziehung von sozialer Schicht zur Befragtenrolle gibt, die ihrerseits SD-Tendenzen unterdrückt.

Das Modell ist - es sei ausdrücklich betont - nur die (auch noch nicht restlos vollständige) - Ausdifferenzierung der insgesamt zu beachtenden theoretischen Variablen zur Erklärung des Befragtenverhaltens. Eine unmittelbare Überführung in ein dann auch empirisch überprüfbares Modell ist ohne weitere Vereinfachungen sicher noch nicht möglich. Immerhin können nun aber zuvor zu undifferenziert betrachtete Vorgänge genauer rekonstruiert werden. So sind z.B. systematische Interviewereffekte als Resultat der situationellen Veränderung der Handlungssituation für den Befragten dann zu erwarten, wenn der Interviewer in besonderem Maße das "situationale looking good" (also: die Richtung), die Relevanz, Transparenz und Isolierung für situationale SD-Tendenzen beeinflußt. Dies ist am ehesten bei "furchterregenden" oder sonstwie wichtig erscheinenden Inhalten, bei klaren Zuordnungen des Interviewers zu bestimmten Öffentlichkeiten und bei (vermuteter) fehlender Konsequenzenisolierung der Fall. Aber auch dann müssen sich Interviewereigenschaften noch keineswegs systematisch auswirken: Bei hoher Determination des AV durch eine bestimmte Einstellung, bei geringen, auf eine Antwortanpassung bezogenen Nutzenerwartungen (z.B. wenn "soziale Anerkennung" nur wenig Nutzen stiftet), bei starken kulturellen SD-Tendenzen und - nicht zuletzt - bei einer starken Orientierung an der Befragtenrolle bleibt für derartige Interviewereffekte kaum Spielraum. Vielleicht mag dieses, zusätzlich zur Komplexität der verschiedenen Aspekte der Interviewereffekte allgemein, erklären, warum solche Effekte weder häufig noch konsistent nachgewiesen werden. Daraus zu schließen, daß die sozialen Prozesse im Interview alle nur aufgabengerecht abliefen, wäre aber sicherlich ebenfalls nicht angebracht.

5. Probleme der empirischen Überprüfung von Interviewereffekten

Die theoretische Analyse hatte ergeben, daß Interviewereffekte die Folge sowohl des Interviewerverhaltens (Fälschung, einstellungsbezogene Vercodung, Befragtenbeeinflussung) wie des Befragtenverhaltens (situationale SD-Tendenz) sein können. Die empirischen Analysen zur Untersuchung von Interviewereffekten gehen demgegenüber i.d.R. so vor, daß Interviewereffekte in Abhängigkeit gewisser äußerer Merkmale (z.B. Alter, Geschlecht, Schichtzugehörigkeit) und/oder bestimmten internen Dispositionen (Einstellungen, Erwartungen) des Interviewers gesehen werden: Verändern sich - ggf. unter Kontrolle anderer Variablen - die Randverteilungen bzw. die erklärten Varianzen bei einer abhängigen Variablen, wenn man gewisse Interviewermerkmale zu den exogenen Variablen hinzufügt, dann liegen Interviewereffekte vor (vgl. den Beitrag von Schanz und Schmidt 1983, in diesem Band, in dem - völlig zu Recht - der multivariate Charakter des Problems ernstgenommen wird). Unbefriedigend ist dabei häufig, daß Interviewereffekte weder regelmäßig, noch gerade immer bei den am "anfälligsten" geltenden Fragen und Situationen festgestellt werden können. Unbefriedigend bleibt dann weiter, daß die Entstehung vorgefundener Effekte notgedrungen immer nur ex post erklärt werden kann. Dieses liegt, so sei vermutet, allerdings weniger an der mangelnden Leistungsfähigkeit der Analyseinstrumente als daran, daß bislang eine deduktive Theorie zur Erklärung von Interviewereffekten fehlt, und daß - nicht zuletzt auch deswegen - in den Erhebungen wichtige Variablen nicht gemessen wurden, die es erlauben würden, nicht nur die "Netto"-Effekte, sondern auch speziellere kausale Prozesse zu untersuchen.

Im folgenden soll - im Anschluß an die theoretische Analyse des Interviewer- und Befragtenverhaltens - versucht werden, in aller gebotenen Vereinfachung ein derartiges deduktives Modell zu entwickeln; hiervon ausgehend soll dann gezeigt werden, unter welchen Bedingungen Interviewereffekte (eines bestimmten Typs) sich überhaupt im Befragungsergebnis zeigen können und welche Variablen zur Analyse der gesamten Vorgänge gemessen werden müßten. Dabei wird sich allerdings zeigen, daß bereits die im ALLBUS 1980 gemessenen Variablen erheblich weitreichendere Aussagen erlauben, als die üblicherweise hierzu erhobenen Merkmale.

Die Problematik bei der empirischen Analyse von Interviewereffekten besteht zunächst aus einem doppelten Theoriedefizit: Erstens muß geklärt werden, über welche genauen kausalen Mechanismen (Interviewerverhalten, Wahrnehmung von Interviewermerkmalen, Befragtenverhalten) ein bestimmter Effekt erklärt werden soll. Und zweitens müßte dann noch geklärt werden, wie die empirisch gemessenen Merkmale (z.B. Alter, Geschlecht, Schichtzugehörigkeit) mit den theoretischen Variablen zusammenhängen: Welche Nutzenfunktionen für welche Zielsituationen mit welchen Erwartungen haben bestimmte Interviewer in bezug auf welchen Typ von Situation und Befragten? Wie werden vom Interviewer mit bestimmten Merkmalen Relevanz, Transparenz und Isolierung einer gegebenen Situation eingeschätzt? Welche Richtung eines "situationalen looking good" signalisieren die sichtbaren Interviewermerkmale bestimmten Befragten usw.?

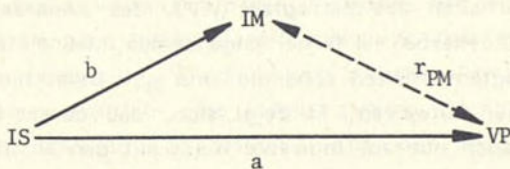
Die Aufzählung zeigt, daß bereits die Verknüpfung der Interviewermerkmale mit den theoretischen Variablen äußerst problematisch ist. Diesem soll hier auch nicht im einzelnen nachgegangen werden: Gewisse Vereinfachungen und Annahmen zu Beginn der Modellbildung (die später schrittweise verändert oder erweitert werden könnten) werden erforderlich. Zur Überführung der o.a. theoretischen Modelle in empirisch überprüfbare Kausalmodelle zur Erklärung von Interviewereffekten seien zunächst die vier o.a. Mechanismen für Interviewereffekte (Fälschungen, einstellungsgemäße Codierungen, Befragtenbeeinflussung, situationale SD-Tendenz der Befragten) in jeweils getrennten Modellen dargestellt. Diese Modelle werden dann zu einem Gesamtmodell zusammengefügt, aus dem sich ersehen läßt, daß Interviewereffekte eines bestimmten Typs nur unter speziellen Bedingungen aus den Interviewermerkmalen allein identifizierbar sind.

In starker Vereinfachung der theroretischen Argumente in Abschnitt 3 sei hier von insgesamt fünf Modellvariablen ausgegangen: Die interne Struktur der Interviewerpersönlichkeit (einschließlich Nutzenfunktionen, Erwartungen, Einstellungen usw.); diese Variable sei IS genannt. Dann die "sichtbaren" Interviewermerkmale (IM). Drittens ein Aspekt des Interviewerverhaltens, der oben mit "Befragtenbeeinflussung" umschrieben worden war (VB). Dann das Befragtenverhalten als Abgabe einer Antwort im Interview (AV). Und schließlich die Protokollierungshandlung des In-

terviewers (VP), die letztlich alleine das auswertbare Interviewergebnis bestimmt. Die folgenden Überlegungen orientieren sich einmal am Grundmodell des Interviewverlaufs (vgl. Schema 1) und dann an den in Abschnitt 3 dargelegten detaillierteren Erklärungen des Interviewer- und Befragtenverhaltens.

In einem ersten Modell seien zunächst Fälschungen und einstellungsbezogene Codierungen gemeinsam erklärt (da beide Handlungsweisen v.a. von der "sonstigen" Situation, insbesondere dem Befragtenverhalten, relativ unabhängig sind). Es wird dabei davon ausgegangen, daß die "interne" Struktur des Interviewers unmittelbar das Protokollierungshandeln beeinflusst. Gleichzeitig wird angenommen, daß die sichtbaren Interviewermerkmale mit der internen Struktur korrelieren bzw. hierfür "Indikatoren" sind. Zunächst werden "externe" Situationsvariablen (außer "Interviewer" und "Befragtem" jeweils) nicht beachtet; d.h. Relevanz, Transparenz und Isolierung - als Resultat der Situationswahrnehmung - werden alleine, ebenso wie die Richtung der Handlungstendenzen, von Interviewer und Befragtem bestimmt. Dann werden Fälschungen und/oder einstellungsgemäße Codierungen unmittelbar von der internen Struktur des Interviewers (Aufgaben- vs. Sozialorientierung und Richtung der Einstellung) bestimmt. Es ergibt sich also das folgende einfache Modell (Schema 5):

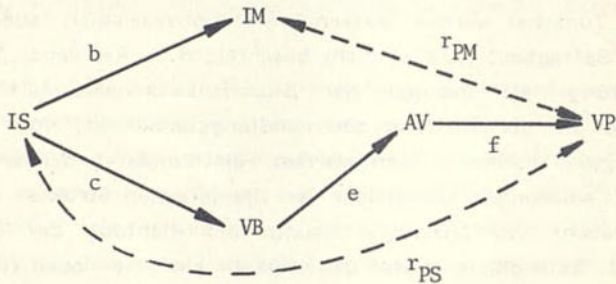
Schema 5: Kausalmodell zur Erklärung von Fälschungen und einstellungsbezogenen Codierungen



Bereits hier wird deutlich, daß bei einer bloßen Erhebung der sichtbaren Interviewermerkmale nur dann auch Interviewereffekte beobachtet werden können (als Korrelation von IM mit VP), wenn IS und IM miteinander zusammenhängen. Die Beziehung zwischen IM und VP ist nur "spurious" ($r_{PM} = ba$; bei der Indizierung der Koeffizienten werden jeweils die zweiten Buchstaben der Variablenkürzel verwendet).

Ein zweites Modell beschreibt die Erklärung von Befragtenbeeinflussungen durch den Interviewer (Schema 6).

Schema 6: Kausalmmodell zur Erklärung von Befragtenbeeinflussung

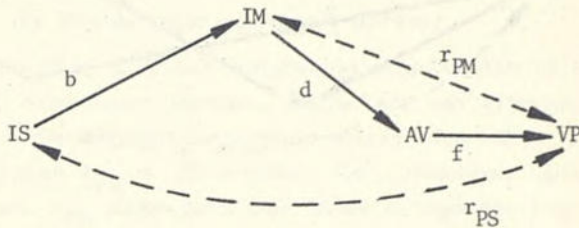


Es wird in diesem Modell angenommen, daß Befragtenbeeinflussungen v.a. das Resultat personaler Interessen des Interviewers sind. Die Beeinflussungshandlungen (VB) bestimmen dann - sicher nicht deterministisch - das Antwortverhalten des Befragten (AV), das seinerseits die Protokollierung bedingt. Hierbei sei ferner angenommen, daß die Protokollierungen nur vom Befragtenverhalten abhängig sind und damit nur einen unsystematischen Fehler aufweisen. Es zeigt sich, daß derart zustande gekommene Effekte auch nur auf indirekte Weise mit den sichtbaren Merkmalen der Interviewer zusammenhängen und erneut nur bei einer relativ engen Korrelation von IS und IM feststellbar wären. Nunmehr ist auch die Beziehung von IS mit VP lediglich indirekt.

Die situationsorientierten Reaktionen von Befragten (in bezug auf den Interviewer) seien schließlich in Schema 7 dargestellt. Danach reagieren die Befragten unmittelbar auf die sichtbaren Interviewermerkmale. Das

hieraus resultierende Befragtenverhalten werde dann entsprechend vom Interviewer protokolliert.

Schema 7: Kausalmodell zur Erklärung des situationsorientierten Befragtenverhaltens

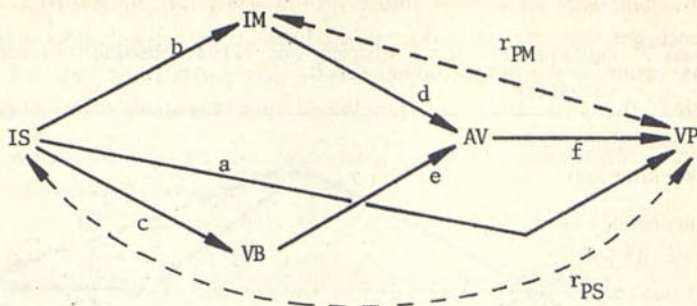


Erneut zeigt sich, daß die - üblicherweise nur untersuchte - Beziehung zwischen den sichtbaren, aber auch den unsichtbaren Interviewermerkmalen und dem Interviewergebnis nur indirekt ist.

Welche forschungspraktischen Konsequenzen ergeben sich nun aus diesen Modellbeziehungen? Dabei ist zu beachten, daß i.d.R. ja Daten über die Interviewer nur unvollständig, geschweige denn Messungen über das Verhalten im Interview vorliegen. Der Regelfall ist, daß nur das Interviewprodukt (VP) verfügbar ist; dann läßt sich natürlich nichts über Interviewereffekte aussagen. Aber auch - und hierauf sei besonders verwiesen - das bloße Erheben von sichtbaren Merkmalen des Interviewers hilft bei der Identifikation von Interviewereffekten nicht viel weiter. Nun kann man zwar die Korrelation r_{PM} als Beziehung zwischen Protokollierung und sichtbarem Merkmal ermitteln (die jeweils zweiten Buchstaben der Variablenkürzel indizieren die Koeffizienten). Aber man kann weder die Abwesenheit von Interviewereffekten bei einem $r_{PM} = 0$ ausschließen, noch identifizieren, welcher Art ein gefundener Effekt ist.

Um dieses zu verdeutlichen, sei zunächst ein Gesamtmodell aller Einzelbeziehungen aufgestellt (Schema 8; die jeweiligen Pfadkoeffizienten wurden mit Buchstaben gekennzeichnet).

Schema 8: Gesamtmodell zur Erklärung von Interviewereffekten



Die verschiedenen Einzeleffekte lassen sich dann über die Pfadkoeffizienten kennzeichnen:

- | | |
|---|-----|
| - Fälschung; einstellungsgemäße Codierung | a |
| - Befragtenbeeinflussung | cef |
| - situationsorientiertes Befragtenverhalten | df |

Für die Korrelation zwischen den sichtbaren Interviewermerkmalen und dem Interviewprodukt r_{PM} und für die Korrelation zwischen den internen Interviewermerkmalen und dem Interviewprodukt r_{PS} ergibt sich dann:

$$(1) \quad r_{PM} = ba + bcef + df$$

$$(2) \quad r_{PS} = a + cef + bdf$$

Man habe nun empirisch ermittelt, daß r_{PM} gleich Null sei. Selbstverständlich könnte dieses Ergebnis die Folge dessen sein, daß gilt: $ba = bcef = df = 0$ (also: daß keinerlei Interviewereffekt vorliegt). Es ist aber leicht ersichtlich, daß das empirische Ergebnis auch die Folge von einander neutralisierenden gegensinnigen Effekten sein kann. Etwa: Wenn ba negativ ist und im Betrag genau den Wert der Summe ($bcef + df$) umfaßt. Es sind zahlreiche - inhaltlich durchaus plausible - Möglichkeiten denkbar, die ein solches "empirisches" Ergebnis der Abwesenheit von Inter-

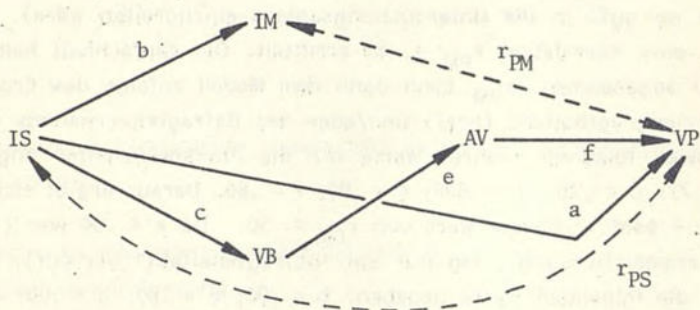
viewereffekten auch bei Vorliegen sehr massiver Hintergrundeffekte erklären können. Die forschungspraktische Relevanz dieses Sachverhaltes liegt darin, daß durch die Veränderung eines Teils der Variablen (z.B. Veränderung der internen Struktur der Interviewer durch ein anderes Schulungs- oder Vergütungssystem so, daß etwa der Koeffizient a gleich Null würde) nunmehr unmittelbar - bei "äußerlich" unveränderter Untersuchungsanlage - deutliche Interviewereffekte festzustellen wären (obwohl "in Wirklichkeit" die Effekte sogar verringert wurden).

Andererseits kann aber auch bei Feststellung von Effekten (d.h.: wenn $r_{PM} \neq 0$) nicht entschieden werden, welche Art von Effekten vorliegt (und wo ggf. in die Untersuchungsanlage einzugreifen wäre). Man habe z.B. eine Korrelation $r_{PM} = .50$ ermittelt. Der Einfachheit halber werde $a = 0$ angenommen. r_{PM} kann dann dem Modell zufolge das Ergebnis des Interviewerverhaltens (bcef) und/oder des Befragtenverhaltens (df) sein. Es seien folgende "wahre" Werte für die Pfadkoeffizienten angenommen: $b = .71$; $c = .40$; $d = .625$; $e = .00$; $f = .80$. Daraus ergibt sich für r_{PM} ($= a + bcef + df$) ein Wert von $r_{PM} = .50$. Da $e = .00$ war (keine Befragtenbeeinflussung), lag nur ein "Befragteneffekt" vor (df). Nun seien aber die folgenden Werte gegeben: $b = .90$; $c = .80$; $d = .00$; $e = .87$; $f = .80$. Hieraus ergibt sich ebenfalls $r_{PM} = .50$. Da $d = .00$ war, lag nunmehr ein bloßer interviewerbezogener Effekt vor. Im Beispiel sind die Koeffizienten, die das empirische Ergebnis erklären, relativ hoch. Wenn man jedoch auch einen direkten Einstellungseffekt a annimmt, wird das Beispiel schon realistischer. Es bleibt festzuhalten: Eine Identifizierung der "Ursachen" festgestellter Effekte ist bei bloßer Erhebung nur sichtbarer Interviewermerkmale ebensowenig möglich, wie ein Ausschluß von Interviewereffekten bei nicht feststellbaren empirischen Beziehungen zwischen Interviewermerkmalen und Interviewprodukt.

Das Bild ändert sich sofort, wenn man auch Messungen der internen Dispositionen der Interviewer hat. Nun verfügt man über drei empirische Beziehungen: r_{PM} , r_{MS} und r_{PS} . In diesem Fall lassen sich Interviewereffekte nicht mehr völlig "neutralisieren": Sofern Effekte überhaupt vorliegen, muß einer der beiden Koeffizienten r_{PM} oder r_{PS} verschieden von Null sein (auch bei Vorliegen gegenläufiger Einzeleffekte). Und zweitens kann man nun grob zwischen interviewerbezogenen und befragtenbezogenen Effekten unterscheiden.

Um dies zu zeigen, sei zunächst wieder von einer Korrelation $r_{PM} = 0$ ausgegangen. Es ist nun weiter davon auszugehen, daß ohne Messungen der Verhaltensvariablen VB und AV man zwischen a und cef (als die beiden zu unterscheidenden interviewerbezogenen Effekte) nicht unterscheiden kann, sondern immer nur den Gesamteffekt ($a + cef$) feststellen kann. Dies kann man Schema 9 entnehmen, in dem nur die interviewerbezogenen Effekte aufgeführt sind.

Schema 9: Kausalmodell der interviewerbezogenen Interviewereffekte



Die empirische Korrelation r_{PS} setzt sich aus dem direkten Effekt a und dem indirekten Effekt cef zusammen. Da aber nur IS und VP gemessen wurden, ist eine Feststellung der Einzeleffekte nicht möglich.

Es ist nun zu zeigen, daß bei einem gefundenen Wert von $r_{PM} = 0$ die Korrelation r_{PS} nur dann auch gleich Null sein kann, wenn df und der gesamte interviewerbezogene Effekt ($a + cef$) beide gleich Null sind (sofern b kleiner als 1 ist). Die Ausgangsfeststellung $r_{PM} = 0$ läßt sich wie folgt zerlegen (bei angenommenem "Ausgleich" der Effekte):

$$\begin{aligned}
 (3) \quad r_{PM} &= (ba + bcef) - df \\
 &= b(a + cef) - df = 0
 \end{aligned}$$

Daraus ergibt sich:

$$(4) \quad df = b(a + cef)$$

Nun kann aber andererseits eine Korrelation $r_{PS} = 0$ (über einen Ausgleich gegenläufiger Effekte) nur dann zustande kommen, wenn weiter gilt:

$$(5) \quad r_{PS} = (a + cef) - bdf = 0$$

Daraus ergibt sich entsprechend:

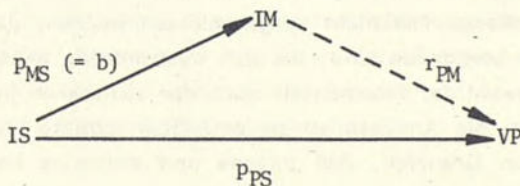
$$bdf = (a + cef)$$

$$(6) \quad df = (a + cef)/b$$

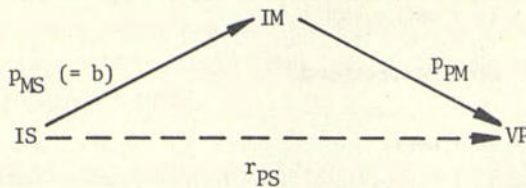
Wenn man nun davon ausgeht, daß b kleiner als 1.0 ist, dann gilt offenkundig $(a + cef)/b \neq b(a + cef)$. Es muß also, sofern nicht $(a + cef)$ und df beide gleich Null und sofern nicht b gleich 1.0 sind, bei $r_{PS} = 0$ gelten, daß $r_{PM} \neq 0$; und bei $r_{PM} = 0$, daß $r_{PS} \neq 0$. Die Differenz zwischen den Werten wird dabei um so größer, je kleiner die Korrelation der internen und externen Interviewermerkmale $r_{MS} = b$ ist. Da man aber wohl davon ausgehen kann, daß sichtbare Merkmale interne Zustände nicht deterministisch anzeigen (außer bei starken kollektiven Identitäten), kann man davon ausgehen, daß "tatsächlich" vorliegende Interviewereffekte auch identifizierbar sind, wenn IS und IM gemessen worden sind.

Um zu zeigen, daß man bei Messung der internen und der externen Interviewermerkmale auch grob zwischen interviewerbezogenen und befragtenbezogenen Effekten unterscheiden kann, seien die interviewerbezogenen und befragtenbezogenen Effekte zunächst in zwei Modellen dargestellt, die nur die gemessenen Variablen enthalten (Schema 10a und b).

Schema 10a: Modell der interviewerbezogenen Effekte (ohne Verhaltensvariablen)



Schema 10b: Modell der befragtenbezogenen Effekte (ohne Verhaltensvariablen)



Bei ausschließlich vorliegenden interviewerbezogenen Effekten (Schema 10a) gilt für die Korrelation $r_{PM} = r_{MS} r_{PS}$ (da hier $p_{MS} = r_{MS}$ und $p_{PS} = r_{PS}$). Der empirische Wert des Pfades p_{PM} muß sich dann als Null ergeben. Ist $p_{PM} \neq 0$, dann müssen (auch) befragtenbezogene Effekte vorliegen (die sich über die Verhaltensvariablen verdeckt mit $p_{PM} = p_{VM} p_{PV}$ vermitteln).

Bei ausschließlich befragtenbezogenen Effekten (Schema 10b) ergibt sich analog, daß dann für die Korrelation $r_{PS} = r_{MS} r_{MP}$ sein muß (da nun $r_{MS} = p_{MS}$ und $r_{PM} = p_{PM}$). Nunmehr muß der Pfad p_{PM} gleich Null sein.

Häufig allerdings werden beide Pfade ungleich Null sein; die Höhe der Pfadkoeffizienten könnte man dann als Hinweis auf die relative Stärke der beiden Effekte nehmen.

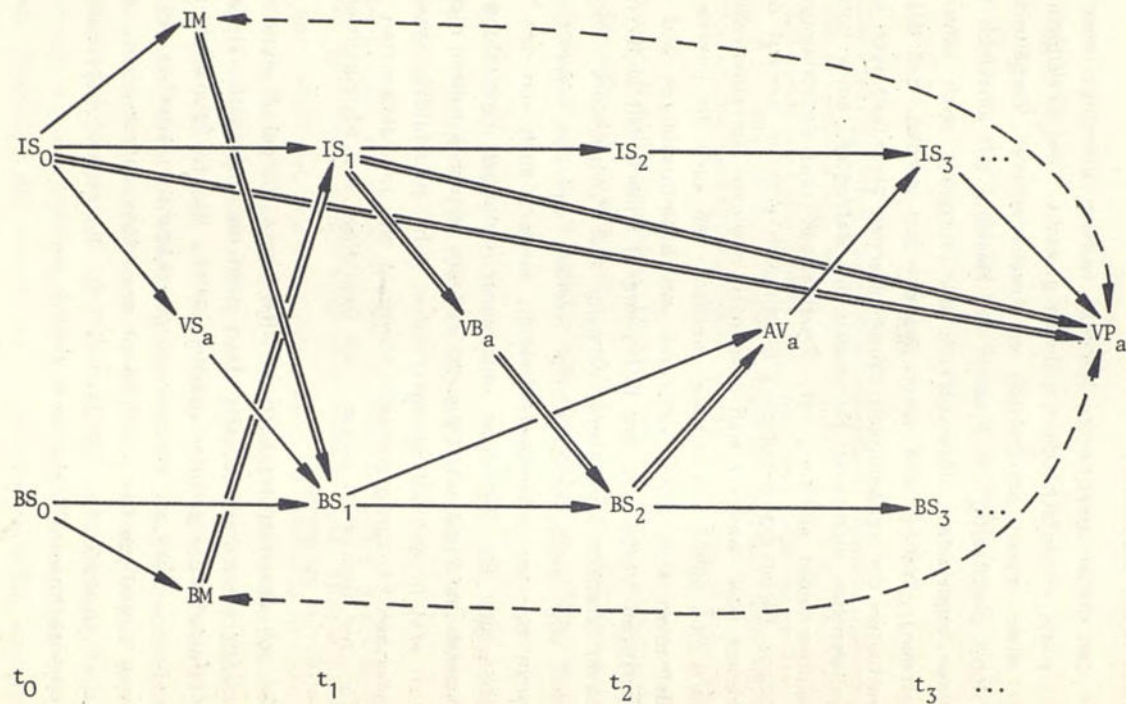
Es sei zusammengefaßt: Bei Messung nur der Protokollierung des Interviewergebnisses läßt sich - selbstverständlich - weder die Anwesenheit noch die Art von Interviewereffekten nachweisen. Bei Messung nur der "sichtbaren" Interviewermerkmale läßt sich zwar zeigen, daß Interviewereffekte überhaupt vorliegen; aber die Art des Effektes kann nicht identifiziert werden. Zudem kann auch bei einem "empirisch" nicht feststellbaren Effekt in diesem Fall nicht ausgeschlossen werden, daß dennoch bestimmte Effekte vorhanden sind, die sich wechselseitig neutralisieren. Erst bei Messung sowohl der internen als auch der sichtbaren Interviewermerkmale kann auch die Anwesenheit gegenläufiger Effekte (bis auf den unwahrscheinlichen Grenzfall, daß interne und sichtbare Merkmale perfekt

korrelieren) festgestellt werden. Allerdings ist dann auch noch nicht die Trennung von Einstellungs- und Beeinflussungseffekten möglich. Hierzu wären Messungen des Verhaltens von Interviewer und Befragtem im Interview selbst notwendig.

In den bisher diskutierten Modellen wurden allerdings immer noch einige, zu stark vereinfachende Annahmen gemacht. Drei Erweiterungen scheinen in einer ersten Annäherung an "realistischere" Verhältnisse notwendig (wobei gleichzeitig das Problem der Messung der einzelnen Variablen und einer empirischen Untersuchung des Vorgangs noch schwieriger lösbar werden): Erstens muß davon ausgegangen werden, daß das Interviewerverhalten (in Abhängigkeit seiner internen Struktur) auch von den wahrgenommenen sichtbaren Merkmalen des Befragten, sowie von dessen Antwortverhalten abhängig ist; Erwartungen und Einstellungen (kaum dagegen Fälschungsneigungen) bilden sich erst im Verlauf der Befragung heraus bzw. wandeln sich auch nach gewissen Befragtenreaktionen. Zweitens muß davon ausgegangen werden, daß auch die interne Struktur des Befragten sich (etwa aufgrund von Beeinflussungen und aufgrund des sonstigen Verhaltens des Interviewers) ändert. Mit anderen Worten: Der zuvor statisch konzipierte Vorgang muß "dynamisiert" werden. Drittens muß aber auch berücksichtigt werden, daß das Antwortverhalten (AV) nicht nur von Interviewereffekten, sondern auch von den "wahren" Einstellungen des Befragten und seiner sonstigen Einbindung in kulturelle Normen abhängig ist; genauer: daß die interne Struktur nicht ausschließlich erst in der Befragungssituation sich herausbildet, sondern von einem gewissen "Ausgangszustand" abhängig ist. Interaktionisten und Artefaktforscher übersehen dieses - die Grundlage des Instrumentes Interview - meist nur zu gerne.

Es soll abschließend noch in aller Kürze ein derart dynamisiertes Modell skizziert werden, für das dann noch die wichtigsten Einzelprozesse und "Effekte" kurz erläutert werden sollen. Das so "dynamisierte" Modell soll dabei - anders als die oben vorgeschlagenen statischen Modelle - weniger eine Grundlage für unmittelbare empirische Forschungen, als ein Versuch einer theoretischen "Entzerrung" der bisher nur untersuchten bzw. nur untersuchbaren "statischen" Effekte sein.

Schema 11: Dynamisiertes und erweitertes Modell zur Erklärung von Befragungsergebnissen (einschließlich von Interviewereffekten)



Das derart dynamisierte Modell knüpft an die Variablen des in Abschnitt 1 dargestellten Grundmodells des Interviewverlaufs an. Hier sind jedoch zusätzlich die wechselseitigen "Wahrnehmungen" auch anderer Situationsmerkmale als nur der Fragestimuli und auch andere Verhaltensweisen als nur die reinen "Fragehandlungen", Fragebeantwortungen und Reaktionscodierungen aufgeführt.

Das Modell beginnt mit einem Zeitpunkt t_0 (Beginn der Befragung), zu dem Interviewer und Befragter gewisse interne Zustände (IS_0 und BS_0) sowie bestimmte sichtbare Merkmale (IM und BM) aufweisen. Die sichtbaren Merkmale werden als "konstant" angenommen; daher gibt es für sie nur einen Wert über die Zeit hinweg (wenngleich man durchaus auch hier Änderungen in den Wahrnehmungen der sichtbaren Merkmale im Interviewverlauf vermuten könnte). Für die internen Zustände wird angenommen, daß sie sich verändern können, und zwar in Abhängigkeit der eingeführten konstanten Situationsmerkmale (IM für den Befragten, BM für den Interviewer) und in Abhängigkeit des variablen Verhaltens der Beteiligten. Diese Veränderung des internen Zustandes ist jedoch nicht nur situationsabhängig, sondern auch von den vorhergehenden internen Zuständen bestimmt; entsprechende Beziehungspfeile kennzeichnen dies. Zu erklären ist das Interviewergebnis in bezug auf eine Frage a ; also: die Codierungshandlung des Interviewers für die Frage a (VP). Der "genetische" Prozeß beginne mit dem Stellen der Frage a durch den Interviewer (VS). Die Frage werde vom Befragten wahrgenommen und führe zu einer Veränderung des internen Zustands zum Zeitpunkt t_1 . Gleichzeitig führe der Interviewer (unter oben geschilderten Umständen) Beeinflussungshandlungen (VB) durch, die den Befragten erneut in seiner internen Struktur ändern (BS_2). Die interne Struktur des Befragten zum Zeitpunkt 2 bestimme nun sein Antwortverhalten in bezug auf Frage a (und alle anderen Einflüsse). Dieses Antwortverhalten werde vom Interviewer (zum Zeitpunkt 3) wahrgenommen und in eine entsprechende Vercodierungshandlung (VP) umgesetzt. Hier wird deutlich, daß eine "korrekte" Codierung auch von der Wahrnehmung der Befragtenreaktion AV abhängig ist; wenn diese nur "diffus" ist, dann bleibt entsprechend mehr Raum für andere Bestimmungsgründe der Codierung. Zu diesen anderen Bestimmungsgründen wären die interviewerbezogenen Interviewereffekte zu zählen: Fälschungen, einstellungsbezogene Codierungen und - sie seien hier noch weiter unterschieden - Erwartungen über rollengemäße Befragtenreaktionen (vgl. Abschnitt 1). Fälschungen und einstellungsbezogene Codierungen sind, den o.a. Ausführungen zufolge, das Resultat des internen Interessenzustands zu Beginn der Befragung. Dieser Effekt wird durch eine direkte Beziehung zwischen IS_0 und VP angedeutet. Im dynamisierten Modell kann jedoch auch berücksichtigt werden, daß diese "Interessen" sich im Verlauf der Befragung ändern; daher müßte man zur Erklärung jeder weiteren Frage (b, c, \dots) jeweils einen "neuen" Ausgangspunkt annehmen, der von dem zuletzt gezeigten Zustand abhängig wäre. Und von diesem "neuen" Ausgangspunkt und den dort vorliegenden "Interessen" müßte dann die entsprechende Fälschungs- bzw. einstellungsbezogene Codierungstendenz erklärt werden. Die Erwartungen des Interviewers über ein "rollengerechtes" Befragtenverhalten konnten in den o.a. Modellen nicht analysiert werden, da dort der Befragte nicht als Situationsbestandteil des Interviewers vorkam. Hier ist jedoch BM aufgeführt, von wo diese Erwartungen her strukturiert werden (z.B. daß langhaarige männliche Befragte liberale Erziehungsziele haben). Dabei ist noch zu beachten, daß diese Erwartungen auch durch das Antwortverhalten

erst strukturiert werden können (z.B. daß eine Person protestantisch ist; und es wird dann die Erwartung aktualisiert: Protestanten trinken nicht). Hier wird davon ausgegangen, daß dieser Einfluß sich nicht mehr auf die Frage a, sondern auf darauf folgende Fragen auswirkt; eine entsprechende Erweiterung des Modells hätte das zu berücksichtigen.

Diese über BM gebildeten Erwartungen haben dann, so die Annahme, einen direkten Effekt auf das Codierungshandeln. In den o.a. statischen Modellen waren Erwartungen deshalb nicht von Fälschungen und einstellungsgemäßen Codierungen unterscheidbar, weil die Erwartungen erst mit IS_1 auftreten, während Fälschungen und einstellungsbezogene Codierungen von der Interessenlage zum Zeitpunkt t_0 bereits abhängig sind.

Die Protokollierung ist aber noch von drei weiteren Einflüssen abhängig: Von den Beeinflussungen des Befragten durch den Interviewer (VB), von der situationalen SD-Tendenz des Befragten und - nicht zu vergessen - von der "wahren" Einstellung des Befragten (bzw. - hier davon nicht weiter zu unterscheiden - von kulturellen SD-Tendenzen und internalisierten Normen).

In bezug auf die Beeinflussungen wird nun auch eine Präzisierung möglich: Ein Interviewer beeinflusst nun einen Befragten erst aufgrund der Wahrnehmung von dessen Merkmalen (und nicht nur aufgrund seiner internen Interessenstruktur). Wenn diese Interessenstruktur relativ stabil ist (z.B. bei einer starken Aufgabenorientierung), dann werden auch die sichtbaren Befragtenmerkmale hier kaum einen Einfluß haben. Für die Veränderung in der Wahrnehmung der sichtbaren Befragtenmerkmale gilt das oben Gesagte entsprechend.

Bezüglich der situationalen SD-Tendenz ergibt sich hier kaum eine Änderung gegenüber dem statischen Modell: Der interne Zustand des Befragten wird durch die Wahrnehmung von Interviewermerkmalen beeinflusst und dieser "wirkt" dann indirekt weiter auf AV.

Für das situationsunabhängige Befragtenverhalten (die Wirkung der "wahren" Einstellung und der kulturellen SD-Normen) wird angenommen, daß von BS_1 , dem internen Zustand des Befragten nach der Wahrnehmung des Fragestimulus, ein direkter Einfluß auf das Antwortverhalten ausgeht; je stärker BS_1 durch BS_0 determiniert ist und je stärker dieser direkte Effekt ist, um so "gültiger" ist das Befragtenverhalten.

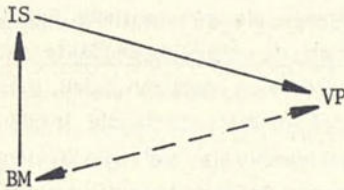
Drei Konsequenzen ergeben sich aus der Dynamisierung der Erklärung von Befragungsergebnissen. Erstens wird nun deutlich, daß die "ideale Befragung" (vgl. Schema 1) ein Spezialfall des vollen Modells ist; die "verzerrenden" Effekte (im Modell mit Doppelpfeilen gekennzeichnet) sind dann gleich Null. Ein "gültiges" Ergebnis ist allerdings auch bei Vorliegen von Verzerrungen möglich, nämlich dann, wenn sich gewisse Effekte neutralisieren.

Zweitens wird in dem vollen Modell auch deutlich, auf welcher 'indirekten' Weise demographische Variablen der Befragten (Alter, Geschlecht, soziale Schicht: BM) mit dem Befragungsergebnis zusammenhängen. Zwar

wird nun ein derartiger Zusammenhang handlungstheoretisch ausdeutbar (die "Pfade" symbolisieren ja "Handlungstendenzen" und Beeinflussungsrichtungen); dieses heißt aber auch, daß man derartige Variablen eigentlich nicht als erklärungsrelevante Größen heranziehen kann. Eine empirische Korrelation von BM mit VP kann nämlich auf sehr verschiedene Weise zustande kommen - nicht nur über eine Beziehung, die über BS und AV unmittelbar verläuft.

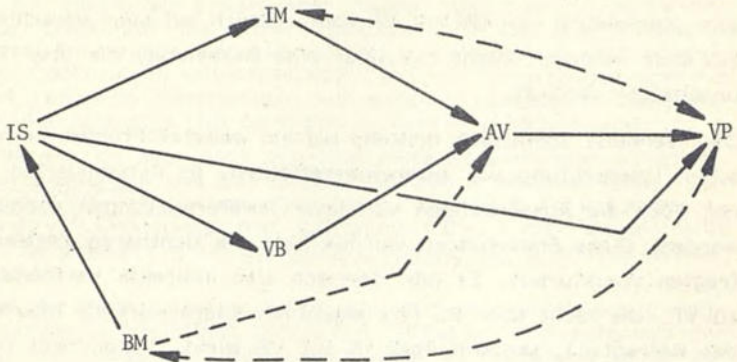
Dies verweist schließlich drittens auf ein weiteres Problem bei der empirischen Überprüfung von Interviewereffekten. Im dynamisierten Modell waren auch die Auswirkungen von Interviewererwartungen analysierbar geworden. Diese Erwartungen wurden über die sichtbaren Merkmale der Befragten strukturiert. Es gibt demnach eine indirekte Verbindung von BM zu VP, die nicht über BS (die eigentlich interessierende interne Struktur des Befragten), sondern über IS auf VP wirkt. Vereinfacht (und erneut statisch) kann man dies so darstellen (Schema 12):

Schema 12: Statisches Modell zur Erklärung von Interviewereffekten über Erwartungen über den Befragten



Wenn man die sichtbaren Befragtenmerkmale einfügt, dann kann nun auch der "substantielle" Bestandteil des Befragtenverhaltens berücksichtigt werden: BM hängt von der (nicht aufgeführten) Größe BS ab, BS - die "interne Struktur" des Befragten - bestimmt AV und damit VP. Die "substantielle" Beziehung zwischen einer sichtbaren demographischen Variablen und einer Befragungscodierung ist also auch im statischen Modell immer nur indirekter Art (vgl. Schema 13; sowie den gebrochenen Pfeil).

Schema 13: Statisches Modell zur Erklärung von Befragungsergebnissen unter Berücksichtigung sichtbarer Befragungsmerkmale



Es wurden zur Verdeutlichung des Problems die Beziehungen zwischen BM und VP (ähnlich wie zwischen IM und VP) als nur über indirekte Beziehungen erklärbar Korrelationen eingetragen. Es wird deutlich, daß die empirische Beziehung zwischen Befragtenmerkmalen und gewissen Befragungsergebnissen nicht ohne weiteres als substantielle Erklärungen der abhängigen Variablen, sondern auch als Interviewereffekte gedeutet werden können. Das Problem verschärft sich dadurch, daß diese indirekte Beziehung zwischen BM und VP nicht nur durch die Interviewererwartungen (nach Maßgabe der Befragtenmerkmale; vgl. die Beziehung BM auf IS, IS auf VP), sondern auch durch Befragtenbeeinflussungen aufgrund von Erwartungen (und Merkmalswahrnehmungen) entstehen kann (vgl. BM auf IS, IS auf VB, VB auf AV, AV auf VP). In einem dynamischen Modell käme sogar noch die situationale SD-Tendenz hinzu, da hier die interne Struktur des Befragten auf der Grundlage der Wahrnehmung von Interviewermerkmalen AV bestimmt und gleichzeitig die interne Struktur mit BM zusammenhängt. Daß die interne Struktur (also nicht nur die sichtbaren "demographischen" Merkmale) auch über Interviewereffekte mit VP zusammenhängen kann, läßt sich aus dem dynamischen Modell ebenfalls leicht ablesen. Im statischen Modell konnten diese Beziehungen aus Vereinfachung

chungsgründen (und wegen der Problematik nicht-rekursiver Modelle) nicht weiter berücksichtigt werden.

Dieses Resultat hat eine wichtige forschungspraktische Konsequenz: Die statistische Überprüfung, wieviel Varianz Befragtenmerkmale bei einer abhängigen Variablen "binden", muß nicht immer auf eine "substantielle" Erklärung hindeuten. Es können sich auch Interviewereffekte dahinter verbergen. Und dieses ist um so wahrscheinlicher, da sowohl Erwartungen unmittelbar als auch Befragtenbeeinflussungen hierbei mitwirken können.

Damit muß aber auch eine Kontrolle "substanzieller" Befragtenmerkmale zur Identifikation von Interviewereffekten die Annahme voraussetzen, daß Erwartungseffekte nicht vorliegen. Da man i.d.R. die zur Überprüfung dieser Annahme notwendigen Erhebungen nicht durchführt, müßte eine entsprechende Annahme dann theoretisch um so besser begründet sein.

Auf eine weitere Kommentierung des Modells sei hier verzichtet. Es dürfte sicher auch so deutlich geworden sein, daß die Untersuchung von Interviewereffekten bzw. die Erklärung des Interviewproduktes eine Vielzahl von Größen zu beachten hat, die - leider - entweder nicht erhoben wurden oder - vielleicht - nicht erhebbar sind. Dies bedeutet konkret: Es müßten nicht nur die sichtbaren Interviewermerkmale, sondern auch Einstellungen, Erwartungen und "Alltagstheorien" über Zusammenhänge von Befragtenmerkmalen und über vermutete Verletzungen von Interaktionsnormen in der Interviewsituation bei den Interviewern erhoben werden (und nicht nur: die Einstellungen zu den Befragungsinhalten). Und es müßten zweitens neben den sichtbaren Befragtenmerkmalen auch deren Erwartungen und Vorstellungen über den Zusammenhang bestimmter Interviewermerkmale und "erwünschten" Reaktionen ermittelt werden. Eigentlich - aber dieses sei hier nur erwähnt und nicht auch ernsthaft vorgeschlagen - müßten auch die Verhaltensweisen im Interview (Stimulusgabe, Beeinflussungen, Antwortgabe) neben dem Protokollierungsergebnis gesondert erfaßt werden. Da alles dieses aber - nicht zuletzt aus Gründen der Forschungsökonomie - jedoch kaum in der erforderlichen Genauigkeit und Gültigkeit auch tatsächlich erhoben werden kann, wird eine theoretische Durchdringung der Prozesse um so mehr erforderlich, damit nicht aus den empirischen Spuren der verdeckten Grundvorgänge die falschen Schlüsse gezogen werden.

Literatur

- Athey, K.R., J.E. Coleman, A.P. Reitman und J. Tang, 1960: Two Experiments Showing the Effect of the Interviewer's Racial Background on Response to Questionnaires Concerning Racial Issues. *Journal of Applied Psychology*, 44: 244-246.
- Atkinson, John W., 1964: *An Introduction to Motivation*. Princeton: D. van Nostrand Co.
- Benney, Mark, David Riesman und Shirley A. Star, 1956: Age and Sex in the Interview. *American Journal of Sociology*, 62: 143-152.
- Blalock, Hubert M. und Paul H. Wilken, 1979: *Intergroup Processes*. New York-London: Wiley.
- Bryant, Eugene C., Isaak Gardner, Jr. und Morton Goldman, 1966: Responses to Racial Attitudes as Affected by Interviewers of Different Ethnic Groups. *The Journal of Social Psychology*, 70: 95-100.
- Cantril, Hadley, 1944: *Gauging Public Opinion*. Princeton: University Press.
- Cosper, Ronald, 1972: Interviewer Effect in a Survey of Drinking Practices. *The Sociological Quarterly*, 13: 228-236.
- Ellis, Albert, 1947: Questionnaire versus Interview Method in the Study of Human Love Relationships. *American Sociological Review*, 12: 541-543.
- Erbslöh, Eberhard und Gerd Wiendieck, 1974: Der Interviewer. S.83-106 in: Jürgen van Koolwijk und Maria Wieken-Mayser (Hrsg.), *Techniken der empirischen Sozialforschung*, Band 4: Erhebungsmethoden, München: Oldenbourg.
- Esser, Hartmut, 1975: *Soziale Regelmäßigkeiten des Befragtenverhaltens*. Meisenheim: Hain.
- Esser, Hartmut, 1982: Forschungsreaktion als soziale Handlung: Zur Erklärung und analytischen Behandlung systematischer Meßfehler bei der reaktiven Datenerhebung. In: Michael Brenner (Hrsg.), *Die Sozialpsychologie sozialwissenschaftlicher Methoden*. Bern: Huber (im Druck).
- Hyman, Herbert H., 1954: *Interviewing in Social Research*. Chicago: University Press.
- Kahn, Robert L. und Charles F. Cannell, 1968: Interviewing: Social Research. S.149-161 in: David L. Sills (Hrsg.), *International Encyclopedia of the Social Sciences*, Band 8. New York: MacMillan and The Free Press.
- Katz, Daniel, 1942: Do Interviewers Bias Polls? *Public Opinion Quarterly*, 6: 248-268.
- Knudsen, Dean D., Hallowell Pope und Donald Irish, 1967: Response Differences to Questions on Sexual Standards: An Interview-Questionnaire Comparison. *Public Opinion Quarterly*, 31: 290-297.
- Robinson, James A., 1960: Survey Interviewing among Members of Congress. *Public Opinion Quarterly*, 24: 127-138.

- Robinson, Duane and Sylvia Rhode, 1946: Two Experiments with an Anti-Semitism Poll. *Journal of Abnormal and Social Psychology*, 41: 136-144.
- Rosenthal, Robert, 1963: On the Social Psychology of the Psychological Experiment: The Experimenter's Hypothesis as an Unintended Determinant of Experimental Results. *American Scientist*, 51: 268-282.
- Schuman, Howard und Jean M. Converse, 1968: The Effects of Black and White Interviewers on Black Responses. *Public Opinion Quarterly*, 35: 44-68.
- Stegmüller, Wolfgang, 1969: Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band 1: Wissenschaftliche Erklärung und Begründung. Berlin-Heidelberg-New York: Springer.
- Williams, J. Allen, Jr., 1964: Interviewer-Respondent Interaction: A Study of Bias in the Information Interview. *Sociometry*, 27: 338-352.
- Wiseman, Frederick, 1972: Methodological Bias in Public Opinion Surveys. *Public Opinion Quarterly*, 36: 105-108.

INTERVIEWSITUATION, INTERVIEWERMERKMALE UND REAKTIONEN VON BEFRAGTEN IM INTERVIEW: EINE MULTIVARIATE ANALYSE

Volker Schanz und Peter Schmidt

1. Einleitung

Diese Arbeit befaßt sich mit Interviewereffekten im Rahmen der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS).

Zunächst wird in einem theoretischen Teil versucht, ein Kausalmodell aus den in der Literatur zum Interview genannten Beziehungen zu formulieren. Um sich vom Ad-hoc-Charakter bisher formulierter Hypothesen und empirischer Ergebnisse zu lösen, wird eine allgemeine Theorie verwendet, mit deren Hilfe Interviewereffekte erklärt werden können. Dabei wird bisweilen exemplarisch - im Vorgriff auf die Analysen - auf eine der abhängigen Variablen ("liberale Erziehungsziele") detaillierter eingegangen.

Danach werden nach spezifizierten (theoretischen und praktischen) Aspekten die abhängigen und unabhängigen Variablen ausgewählt und eine Reihe linearer Modelle aufgestellt.

Das Gewicht von Interviewer-/Interviewmerkmalen wird über einen Vergleich der erklärten Varianz vor und nach der Einführung dieser Merkmale und über die Analyse der Stabilität der Koeffizienten ermittelt. Darüber hinaus werden die Gewichte bestimmter spezifischer Interviewer-/Interviewmerkmale im Detail betrachtet.

2. Problemstellung, Stand der Forschung

In der methodologischen Literatur und in Lehrbüchern der empirischen Sozialforschung (vgl. etwa Atteslander und Kneubühler 1975; Esser 1975; Sudman und Bradburn 1974 sowie Esser 1983 in diesem Band) wird nahezu immer davon ausgegangen, daß das Antwortverhalten eines Befragten nicht ausschließlich vom eigentlichen (intendierten) Fragestimulus bestimmt

wird. Andere Einflußfaktoren wie etwa die "Tendenz zu sozial erwünschten Antworten" und die "Ja-Sage-Tendenz", die Motivation des Befragten sowie die konkrete Interviewsituation wurden in einer großen Anzahl methodologischer Arbeiten nachgewiesen. Diese empirischen Ergebnisse wurden auch zum Anlaß genommen, schwerwiegende Zweifel an der methodologischen Qualität des Interviews zu formulieren. In der gegenwärtigen Forschungspraxis werden die obengenannten, oft auch als "Störvariablen" bezeichneten Einflußgrößen sehr selten zusätzlich gemessen oder bei der empirischen Testung der Hypothesen kontrolliert. Man geht vielmehr beim Testen von Hypothesen bzw. Theorien in der Regel von einem - wie Lakatos (1970) es nennt - "naiven Falsifikationismus" aus. Demonstrieren wir das an einem einfachen Beispiel. Ausgangspunkt sei folgende Hypothese:

H_1 : Je höher der sozioökonomische Status einer Person ist, desto stärker vertritt diese liberale Erziehungsziele.

Bei Widerlegung der Hypothese kann nicht entschieden werden, ob die eigentliche Theorie oder die Hintergrundtheorie in Form der obengenannten Response-Sets und Interviewereinflüsse falsch sind. Daher muß eine theoretische Entscheidung getroffen werden, ob die eigentliche Theorie bzw. Hypothese oder die Hintergrundtheorie modifiziert werden soll. Im Sinne eines möglichst strengen Tests von Theorien müßten solche alternativen Erklärungen durch Einführung und Konstanthaltung der entsprechenden Variablen in einer multivariaten Analyse geprüft werden.

In der klassischen Auffassung bleibt die Hintergrund- bzw. Beobachtungstheorie implizit. Die Wahrheit der Hintergrundtheorie wird a priori vorausgesetzt statt empirisch überprüft. Wir wählen hier statt dessen die Strategie einer strengeren Überprüfung der eigentlichen Theorie, indem eine Teilmenge von Hintergrundannahmen (Interviewereinflüsse) als Hypothesen reformuliert und empirisch geprüft werden.

Interviewereinflüsse können bezüglich des Antwortverhaltens sowohl Verzerrungen der Antworten als auch Antwort- oder Interviewverweigerung bewirken; uns geht es im folgenden um den erstgenannten Aspekt.

Empirisch kann der Begriff "Interviewereinflüsse" (oder: Interviewereffekte) folgendermaßen expliziert werden:

D_0 : Interviewereffekte/-einflüsse soll heißen: diejenige Varianz in den Ergebnissen, die durch Interviewermerkmale (und deren Interaktion mit Befragtenmerkmalen) erklärt werden kann.

Diese Explikation wird im empirischen Teil durch die Anwendung multipler Regressionsanalysen umgesetzt. Dort wird u.a. geprüft werden, ob die erklärte Varianz der jeweiligen abhängigen Variablen durch die Hinzunahme von Interviewermerkmalen steigt oder gleichbleibt.

Gehen wir als nächstes auf den Stand der gegenwärtigen Forschung ein.

2.1 Theoriedefizit

Schon ein kursorischer Überblick über den Stand der Literatur zeigt den ausgeprägten Ad-hoc-Charakter der gegenwärtigen Forschung. Eine ganze Reihe von Arbeiten untersuchen (eindimensional) den Einfluß eines bestimmten Merkmals (etwa: Alter o d e r Geschlecht) des Interviewers auf das Befragtenverhalten.

Beispiele hierfür sind:

für "Alter des Interviewers":

Erbslöh 1973; Sudman und Bradburn 1974; Freeman und Butler 1976;

für "Geschlecht des Interviewers":

Johnson und Delamater 1976; Freeman und Butler 1976;

für "Schulbildung des Interviewers":

Sudman und Bradburn 1974;

für "Status des Interviewers";

Feldman, Hyman und Hart 1951; Sudman und Bradburn 1974; Dijkstra und Elsinga 1979;

zur "Interviewsituation"¹:

Taietz 1962; Pfeil und Friedrichs 1965; Lutynska 1969; Berekhoven et al. 1975; Atteslander und Kneubühler 1975.

Sehr bald allerdings wurde (zum Teil jedenfalls) diese isolierte Betrachtungsweise aufgegeben und eine Interaktion zwischen bestimmten Interviewermerkmalen (etwa: Alter u n d Geschlecht des Interviewers) angenommen und dies mittels einfacher Kreuztabellierung überprüft. Exemplarisch können hier Benney, Riesmann und Star (1956) genannt werden.

Ein entscheidender Schritt wurde unternommen, als man nicht nur eine Interaktion zwischen verschiedenen Interviewermerkmalen untereinander, sondern auch deren Interaktion mit den entsprechenden Befragtenmerkmalen annahm. Folgende Literatur kann hier als Beispiel genannt werden:

zur Interaktion von Alter, Geschlecht und Schulbildung des Interviewers mit den entsprechenden Merkmalen des Befragten:

(teilweise) Hyman 1954: 164; Sudman und Bradburn 1974;

zur Interaktion des Status von Interviewern und Befragten:

eine Reihe von Arbeiten, die konkurrierende Hypothesen zur sozialen Distanz - gemessen als Differenz zwischen den Statusmerkmalen von Interviewern und Befragten - spezifizieren: Williams 1964; Anger 1969; Dohrenwend 1969; Dohrenwend, Colombotes und Dohrenwend 1968; Weiss 1975; Freeman und Butler 1976.

Schon früh wurde zudem diskutiert, daß auch unsichtbare Merkmale der Interviewer als Determinanten des Befragtenverhaltens anzusehen seien. Einige Beispiele hierfür sind:

zu Einstellungen des Interviewers:

Blankenship 1940; Stember und Hyman 1949; Hyman 1954; Cosper 1969; Freitag und Barry 1974; Berekhoven et al. 1975;

zu Wahrscheinlichkeitserwartungen des Interviewers:

Stember und Hyman 1949; Smith und Hyman 1950; Erbslöh 1973; Sudman, Bradburn, Blair und Stocking 1977.

Neben den genannten Faktoren² wurden in der Literatur noch weitere Differenzierungen angedeutet, bislang aber nicht in einem komplexeren Modell überprüft. Generell fehlt jedoch der Versuch, durch Anwendung allgemeiner Theorien die sich widersprechenden Einzelbefunde in der großen Zahl empirischer Ad-hoc-Untersuchungen zu erklären und zu integrieren. Einen Versuch in dieser Richtung stellen die Arbeiten von Sudman und Bradburn (1974), Atteslander und Kneubühler (1975) und Esser (1975) dar. (Eine explizite Verwendung der Wert-Erwartungstheorie findet sich bei Esser (1983, in diesem Band)).

Damit wollen wir die Diskussion des jetzigen Standes der Literatur über Interviewereffekte abschließen und kurz auf methodische Defizite der jetzigen Forschung eingehen.

2.2 Methodendefizit

Bislang überwiegen innerhalb der Forschungen zu Interviewereffekten Methoden wie Mittelwertsvergleiche bestimmter Fragen über Interviewer-subgruppen sowie einfache und mehrdimensionale Kreuztabellen. Gegen diese methodische Vorgehensweise lassen sich aber eine Reihe von Argumenten vorbringen (vgl. Hirschi und Selvin 1967; Hummell und Ziegler 1976, besonders E15-E17):

- (1) Es fehlt bei der Analyse von Tabellen ein Maß, das dem multiplen Determinationskoeffizienten (R^2) vergleichbar ist. Daher kann man an keinem der Punkte der Analyse sagen, wieviel der Varianz der abhängigen Variablen erklärt ist bzw. in welchem Ausmaß die erklärte Varianz zunimmt, wenn zusätzliche unabhängige Variablen (Interviewermerkmale) betrachtet werden.
- (2) Mit dem in (1) dargestellten Problem hängt zusammen, daß man nie genau weiß, ob zusätzlich eingeführte Variablen tatsächlich noch nicht erklärte Varianz erklären.
- (3) Da die unabhängigen Variablen nahezu immer untereinander korreliert sind, sind Schlußfolgerungen über ihre Beiträge äußerst ungenau.
- (4) Da bei detaillierterer Aufgliederung der Untersuchungsgesamtheit die Zahl der Fälle in den einzelnen Zellen sehr schnell und stark schrumpft, kommt man praktisch - unter Berücksichtigung der Fallzahl und der Zahl und Ausprägungen der Variablen - selten zu einer simultanen Betrachtung aller relevanten Variablen.

Mit den genannten Verfahren können also komplexere Ursache-Wirkungs-Modelle, wie sie dem erreichten theoretischen Stand entsprechen, nicht überprüft werden. Daher scheint uns die Anwendung multivariater statistischer Verfahren (z.B. Regressionsanalysen) eine notwendige Alternative zu der bisher praktizierten methodischen Vorgehensweise zu sein.

3. Ein integrativer Vorschlag

3.1 Anwendung der kognitiv-hedonistischen Theorie

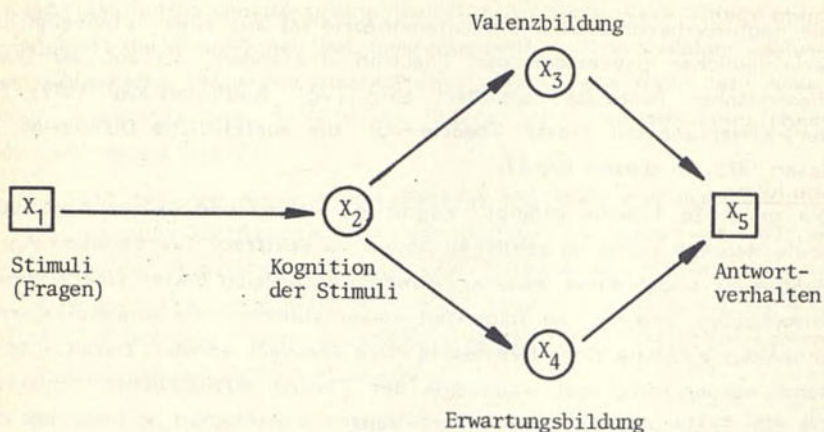
Die kognitiv-hedonistische Verhaltenstheorie ist aus einer Integration unterschiedlicher psychologischer Theorien entstanden, die auf der Ebene theoretischer Postulate formuliert sind (vgl. Kaufmann-Mall 1977; zum Wert-Erwartungsteil dieser Theorie vgl. die ausführliche Diskussion bei Esser 1983, in diesem Band).

Die genannte Theorie erlaubt, Kognition, Valenz und Erwartungsbildung sowie Handeln selbst im zeitlichen Ablauf zu erklären. Die Verwendung als Meßtheorie bildet einen weiteren Anwendungsbereich dieser sehr allgemein formulierten Theorie. Im folgenden sollen zunächst die Anwendungsmöglichkeiten im Sinne einer Meßtheorie kurz skizziert werden. Daran anschließend werden diejenigen Aussagen der Theorie ausführlicher behandelt, die zur Erklärung des Antwortverhaltens von Befragten im Interview dienen können. Am Ende gehen wir dann auf die Probleme ein, die bei der Anwendung auftreten.

Zur anschaulichen Darstellung der Zusammenhänge, die beim Meßprozeß theoretisch ablaufen, verwenden wir Abbildung 1 (S.78); diese Abbildung soll lediglich eine Übersicht ermöglichen und noch kein präzises Strukturgleichungsmodell darstellen. Man sieht, daß Fragen (Stimuli) zunächst einmal kogniziert (wahrgenommen) werden müssen. Die Art der Wahrnehmung, die Art des Dekodierungsprozesses und Fragen der Konzeptbildung dürfte von der Formulierung der Fragen und von Situationsvariablen abhängen.

Der zur Erklärung des Befragtenverhaltens relevante Teil der kognitiv-hedonistischen Theorie wäre der über die Entstehung von Kognitionen. In unserem Zusammenhang wollen wir aber hierauf nicht näher eingehen. Den kognizierten Konzepten werden nun in einem zweiten Schritt Erwartungen und Valenzen zugeordnet. Diese beziehen sich aber nicht nur auf das zu messende theoretische Konstrukt. Interviewermerkmale, Selbstbild, Ausmaß sozialer Wünschbarkeit und Situationsvariablen werden ebenso das Ausmaß der Erwartung von Konsequenzen des gezeigten Antwortverhaltens sowie deren Valenz beeinflussen (vgl. Sudman und Bradburn 1974; Atteslander und Kneubühler 1975; Esser 1975).

Abbildung 1: Ablauf des Messprozesses



Esser (1983, in diesem Band) weist daraufhin, daß die Explananda des Interviewprozesses in drei Einzelaspekte zerfallen:

- (1) Die Reaktion des Befragten in Abhängigkeit von Stimulierungen durch den Interviewer und gewissen eigenen Merkmalen.
- (2) Die Stimulierungen und Reaktionen des Interviewers in Abhängigkeit von Fragebogen und Befragtenreaktionen, weiterer Situation (wozu vor allem die beobachtbaren Merkmale des Befragten gehören) und gewissen eigenen Merkmalen (z.B. Einstellungen zum Fragegegenstand).
- (3) Die Vercodungshandlungen des Interviewers ebenfalls in Abhängigkeit von Befragtenreaktionen, weiterer Situation und eigenen Merkmalen.

Mit den uns zur Verfügung stehenden Daten können aber die drei Aspekte nicht gleichzeitig untersucht werden. Direkte Messungen liegen nur für die Vercodungshandlungen des Interviewers vor. Diese bilden die Rohdaten unserer Hauptbefragung mit 2955 Personen (Hauptstudiendatensatz).

Parallel dazu wurde ein Teil der gleichen Variablen bei den Interviewern erfragt (Eigeninterviewdatensatz).

Wir können somit nicht die kausale Wirkung der vorher genannten Explananda getrennt erfassen. Hierfür wären Prozeßdaten zum Ablauf des Interviews notwendig. Zur Veranschaulichung des Prozeßcharakters sei an dieser Stelle eine Modifikation des Schemas von Esser (1983, in diesem Band) angeführt.

Abbildung 2: Modell des Interviewprozesses



I_0, I_1 = Interviewer zum Zeitpunkt 0, 1

B_1 = Befragter zum Zeitpunkt 1

C_1 = Codierung zum Zeitpunkt 1

Gehen wir nun näher auf den Teil der kognitiv-hedonistischen Theorie ein, der zur Erklärung des Antwortverhaltens herangezogen werden kann. Da Antwortverhalten eine Teilmenge von Handlungen bildet, kann das entsprechende Kausalmodell zur Erklärung von Handlungen auf Antwortverhalten angewendet werden (zur Definition der verwendeten Begriffe vgl. Kaufmann-Mall 1977: 18).

Das theoretische Postulat zur Erklärung von Handlung lautet nun (vgl. Kaufmann-Mall 1977: 24):

TP₁: Die Wahrscheinlichkeit einer Handlung steigt mit dem Produkt aus den Erwartungen (x_1) der Konsequenzen dieser Handlung, gezeigt in der kognizierten Situation (x_2) und den Valenzen dieser Konsequenzen (x_3).

Mit diesem theoretischen Postulat können wir allerdings noch nicht erklären, welchen (exakten) Wert ein Befragter auf einer Skala ankreuzt. Hierzu wird folgende entscheidungstheoretische Zusatzhypothese benötigt (vgl. Kaufmann-Mall 1977, ebenda): Im konkreten Fall wird das Verhalten gezeigt, für das folgendes Produkt ein Maximum ergibt:

$$x_1 * x_2 * x_3 = \max \text{ Handlung}_i$$

Damit kommt zum Ausdruck, daß das Produkt für verschiedene Entscheidungsalternativen, d.h. verschiedene Skalenpunkte, unterschiedlich hoch sein kann und beim Antwortverhalten der Skalenwert angekreuzt wird, für den das Produkt den höchsten Wert annimmt.

Aus dem theoretischen Postulat ergibt sich, daß z.B. das Antwortverhalten dann "verzerrt" ist, wenn eine Frage vom Befragten völlig mißverstanden wird. In der Terminologie des theoretischen Postulats ist diese Frage dann falsch kogniziert. Wenn andererseits die Folgen einer ehrlichen Antwort mit hohen negativen Sanktionen assoziiert werden, wird die gegebene Antwort nicht den "wahren" Wert der zu messenden Variablen repräsentieren. Darüber hinaus läßt sich z.B. auch das Nichtausfüllen eines Fragebogens mit diesem Postulat erklären. Falls einer Person der ganze Fragebogen irrelevant erscheint und damit das Ausfüllen des Fragebogens keinen Wert für sie darstellt, wird das ganze Produkt "Null", und dies impliziert nach dem oben genannten Postulat das Nichtausfüllen des Fragebogens. Die gleiche Konsequenz kann aber auch ein kognitiver Konflikt beim Befragten durch "cross-pressure" haben.

In der hier dargestellten Form bildet die Anwendung dieses theoretischen Postulats aber noch nicht mehr als eine heuristische Orientierung. Konkrete Hypothesen über die Wirkungen der Eigenschaften des Stimulusmaterials, Interviewereigenschaften und der sozialen Situation, in der das Interview durchgeführt wird, müssen noch mit der allgemeinen kognitiv-hedonistischen Theorie konfrontiert werden (vgl. hierzu auch Schema 4 bei Esser 1983, in diesem Band).

3.2 Zuordnung zu dem empirisch geprüften Modell

Das eben skizzierte Modell ist auf der theoretischen, d.h. Konstruktebene formuliert. Um es an Daten zu prüfen, benötigt man empirische Indikatoren für die Valenz der Folgen sowie für Erwartungen. Ebenso benötigt man Informationen darüber, welche Folgen in einer gegebenen Interviewsituation überhaupt relevant sind. Diese Informationen wurden im Rahmen der Allgemeinen Bevölkerungsumfrage nicht erhoben; erfaßt wurden demographische und ausgewählte Verhaltens- und Einstellungsmerkmale der Interviewer. Es muß nun die Verknüpfung dieser Interviewermerkmale mit

dem Vokabular der kognitiv-hedonistischen Theorie durch Explikation entsprechender Hypothesen erreicht werden. Dies wird im folgenden am Beispiel des Konstrukts "liberale Erziehungsziele" versucht.

Wir gehen - wie angedeutet - davon aus, daß das Antwortverhalten des Befragten auch von der Folge (Bestätigung bzw. Anerkennung durch den Interviewer) beeinflußt wird. Aufgrund sichtbarer Merkmale wie Kleidung, Auftreten, Reaktionen bei vorherigen Fragen usw. dürfte der Befragte sich ein Bild, d.h. Hypothesen über den Interviewer gemacht haben, wenn die Frage nach den Erziehungszielen gestellt wird. Es ist nun anzunehmen, daß der Befragte um so mehr Anerkennung und Bestätigung vom Interviewer erwartet, je mehr Ähnlichkeit er mit der von ihm wahrgenommenen Einstellung des Interviewers äußert (vgl. Dijkstra o.J.).

Unsere Erklärungsskizze kann - exemplarisch für das zu erklärende Konstrukt "liberale Erziehungsziele" des Befragten - durch das Pfaddiagramm in Abbildung 3 (S.82) anschaulich dargestellt werden.

Da wir x_3 , x_4 und x_5 nicht empirisch erhoben haben, können wir nur den indirekten Effekt von x_1 , x_2 auf x_7 empirisch fassen. Es handelt sich - technisch gesprochen - um den Koeffizienten der reduzierten Form (Duncan 1975).

Im weiteren gehen wir nun konkret auf die Überprüfung von Interviewer-Effekten in der Allgemeinen Bevölkerungsumfrage ein.

4. Abhängige und unabhängige Variablen

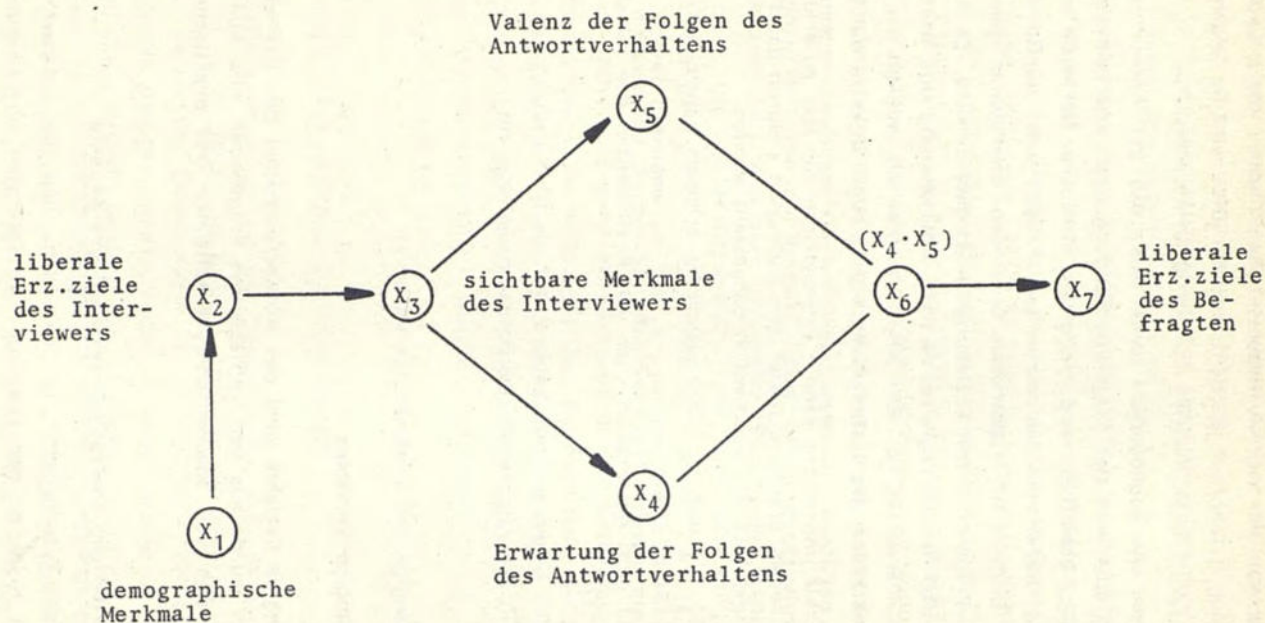
4.1 Abhängige Variablen

Als abhängige Variable wird das Antwortverhalten des Befragten angesehen. Dazu wurde aus dem vorliegenden Fragebogen eine Auswahl getroffen, die sich von theoretisch-systematischen und praktischen Aspekten leiten ließ.

Auswahlkriterium: theoretisch-systematische Aspekte

"Theoretisch-systematisch" soll bedeuten, daß die ausgewählten Fragetypen mit Kriterien der Literatur zum Interview übereinstimmen sollten.

Abbildung 3: Modell 1 - Determinanten der "Erziehungsziele" des Befragten



Folgende Frage-Dimensionen werden in der Literatur hierzu genannt:

- (1) Einstellungsfragen sind eher oder stärker anfällig für Interviewereffekte als Verhaltensfragen (Koolwijk 1969; Ferber und Wales 1972; Sudman und Bradburn 1974; Bradburn et al. 1978);
- (2) heikle bzw. unangenehme Fragen sind eher (stärker) biasanfällig als nicht-heikle Fragen (Koolwijk 1969);
- (3) schwierige Fragen sind eher (stärker) biasanfällig als leichte Fragen (unter der Annahme, daß schwierige Fragen des öfteren ein Eingreifen des Interviewers erfordern und damit interviewerspezifische Bedeutungssetzungen implizieren);
- (4) offene Fragen sind eher (stärker) als geschlossene Fragen biasanfällig (unter der Annahme, daß bei offenen Fragen mehr Raum für Interviewereffekte ist; vgl. dazu Klingemann 1978 und ganz unspezifisch Anger 1969: 601);
- (5) durch den Interviewer nicht überprüfbare Antworten sind eher (stärker) biasanfällig als durch ihn überprüfbare Antworten (unter der Annahme, daß der Befragte bei wenig überprüfbaren Fragen sich der vermuteten Einstellung des Interviewers annähern kann, ohne negative Sanktionen befürchten zu müssen).

Offene Fragen (4) entfallen für diese Arbeit, da in der Allgemeinen Bevölkerungsumfrage (außer zur Erfassung der Berufe) nur mit geschlossenen Fragen gearbeitet wird. Die Dimensionen (2), (3) und (5) sind darüber hinaus nicht trennscharf und nicht unabhängig - d.h. beispielsweise: eine Frage kann sowohl heikel als auch schwierig³ und überprüfbar oder heikel, leicht und überprüfbar (usw.) sein. Was die Differenzierung von Einstellungs- und Fakten- bzw. Verhaltensfragen anbetrifft, so handelt es sich dabei wohl eher um ein Kontinuum als um eine Dichotomie - mit anderen Worten: Fragen haben (bis auf wenige "harte Faktenfragen" wie beispielsweise die Frage nach dem Beruf) mehr oder weniger Einstellungs- oder Faktenbezug. Insofern wir nach einer Aufgabe dieser Differenzierung nach wie vor Fragen mit mehr oder weniger engem Bezug zu Einstellungen und Verhalten haben, hat dies für die Fragestellung zunächst keine Konsequenz.

Auswahlkriterium: praktische Aspekte

Um Interviewereffekte analysieren zu können, wurden eine Reihe von Angaben über die Interviewer erhoben:

- (1) demographische Merkmale wie beispielsweise: Geschlecht, Alter, Religionszugehörigkeit, Kirchengangshäufigkeit, Schulabschluß, Erwerbstätigkeit, berufliche Stellung, gewerkschaftliche Organisiertheit;
- (2) allgemeine Interviewermerkmale (mit Bezug zur Interviewertätigkeit): Anzahl der 1979 durchgeführten Interviews; Zeitverfügbarkeit;
- (3) Einstellungs- und Verhaltensmerkmale der Interviewer: zu Erziehungszielen, politischen Problemen, "oben-unten"-Selbsteinstufung, Parteien/Politiker-Thermometer.

Unser praktisches Problem besteht nun darin: Insoweit in unseren Modellen Verhaltens- oder Einstellungsmerkmale der Interviewer (als Parallelmessungen) von Bedeutung sind, müssen wir unsere abhängigen Variablen zusätzlich noch aus den Bereichen auswählen, zu denen uns Messungen der entsprechenden Interviewermerkmale vorliegen.

Ausgewählte abhängige Variablen

Unter Berücksichtigung der genannten Aspekte haben wir die folgenden fünf Fragen zur Analyse ausgewählt:

Frage 8: liberale Erziehungsziele

Index: Summenscore der Items "Selbständigkeit", "Selbstvertrauen", "Verantwortungsbewußtsein", "Kritikfähigkeit", "Verständnis"⁴.

Je höher der Score, desto stärker betont der Befragte liberale Erziehungsziele.

Frage S46: Oben-Unten-Einstufung

Je höher der Wert der Variablen, desto weiter "oben" ordnet sich der Befragte ein.

Frage 22: subjektive Schichteinstufung⁵

Je höher der Wert der Variablen, desto höher im Schichtgefüge ordnet sich der Befragte ein.

Frage 33: Zweitstimme

dichotomisiert: (1) CDU/CSU versus (2) SPD

Frage 34: Parteisympathie-Score

Index: Differenzscore zwischen SPD- und CDU-Sympathie. Je höher der Score, um so mehr überwiegt die Sympathie des Befragten für die SPD.

Frage 28: Post-Materialismus

Index: (1) Materialisten, (2-3) Mischtypen, (4) Post-Materialisten.

(Basierend auf einer Kombination der Antworten zu den Statements A-D.)

4.2 Unabhängige Variablen

Im folgenden werden zwei Gruppen von unabhängigen Variablen unterschieden:

- (1) "substantielle unabhängige Variablen": Variablen, die in der Literatur für das betreffende Merkmal als solche genannt werden;
- (2) Interviewer-/Interviewmerkmale.

Substantielle unabhängige Variablen

Die sechs abhängigen Variablen lassen sich in vier Gruppen zusammenfassen, für die die folgenden unabhängigen Variablen⁶ spezifiziert werden:

Erziehungsziele: Bargel (1979) und Sandberger und Bargel (1980) verweisen zur Erklärung unterschiedlicher Erziehungsziele oder zur unterschiedlichen Bedeutung, die Befragte Erziehungszielen zuweisen, auf die üblichen sozio-demographischen Standardvariablen (wie: berufliche Stellung, Einkommen, Bildung, Alter). Wir fügen dem das Geschlecht des Befragten als Variable hinzu, weil wir annehmen, daß dies aufgrund des Sozialisationshintergrundes einen Effekt haben könnte. Da wir "liberale" Erziehungsziele erklären wollen, vermuten wir, daß hier Religionszugehörigkeit und Kirchgang einen weiteren Teil der Varianz aufklären werden. Die Variablen subjektive Schichtestufung und den Post-Materialismus-Score nehmen wir gleichfalls als erklärende Faktoren mit auf. (Personen mit ho-

hem Post-Materialismus-Score und höherer Schichteinstufung müßten wohl stärker liberale Erziehungsziele vertreten.)

Subjektive Schichteinstufung; Oben-Unten-Einstufung: Entsprechend Hodge und Treiman (1968), Jackman und Jackman (1973), Klügel et al. (1977), Jöreskog und Sörbom (1981) und Klingemann (1981) sind die zentralen Erklärungsfaktoren sowohl für objektive als auch subjektive Schichtzugehörigkeit berufliche Stellung, Schulbildung und Einkommen - als Orientierungshypothese formuliert: Je höher die objektiven Statusmerkmale, desto höher die subjektive Schichteinstufung. Diesen traditionellen Statusmerkmalen fügen wir eine Dimension "Eigentum" hinzu, die für das subjektive Bewußtsein der Stellung im sozialen Gefüge möglicherweise von Bedeutung sein könnte. Überprüft wird weiterhin der Faktor "Gewerkschaftsmitgliedschaft", da dies durch Jackman und Jackman (1973) nahegelegt wird.

Wir betrachten beide abhängigen Variablen zunächst als Parallelmessungen, für die wir denselben Set an unabhängigen Variablen postulieren.

Zweitstimme; Parteisympathie: Folgt man Küchler (1977), Pappi (1977), Klingemann und Taylor (1977) und Kaltefleiter und Nißen (1980), so läßt sich die Wahlabsicht erklären durch berufliche Stellung, Alter, Geschlecht, Religionszugehörigkeit/Kirchgangshäufigkeit, Gewerkschaftsmitgliedschaft, Parteiorientierung und die Links-Rechts-Einstufung.

Uns interessiert insbesondere die Neigung der Befragten zur CDU, da diese in der Regel in den Wahlrohdaten aus Umfragen unterrepräsentiert ist - vgl. beispielsweise Radtke (1977). Dies könnte unter anderem mit einem Einfluß des Interviewers zusammenhängen.

Wir nehmen an, daß beide abhängige Variablen im wesentlichen von den gleichen unabhängigen Variablen beeinflußt werden (vgl. Berger 1977; Falter 1977).

Post-Materialismus: Nach Inglehart (1979) und Küchler (1983, in diesem Band) sind die zentralen Erklärungsfaktoren für den Post-Materialismus Alter, Beruf, Schulbildung, Konfession und Geschlecht. Wir fügen dem Erwerbstätigkeit und Kirchgangshäufigkeit hinzu.

Interviewer- und Interviewmerkmale:

Hier sollen vier Typen von Interviewer-/Interviewmerkmalen unterschieden werden:

- (1) soziodemographische Merkmale des Interviewers, wie sie in der Literatur über Interviewereffekte genannt werden⁷. Dazu gibt es - wie unter 2.1 Theoriedefizit angedeutet - eine Reihe von Ad-hoc-Hypothesen, es fehlt aber eine intervenierende Theorie, die angibt,
- (a) unter welchen Bedingungen diese Merkmale (wie Alter, Geschlecht des Interviewers) überhaupt einen Einfluß haben;
 - (b) wie dann (gegebenenfalls) die Richtung der Effekte sein wird.

Die Antwort auf beide Fragen dürfte aus einem Verweis auf bestimmte Arten/Qualitäten von Fragen bestehen. Für hier relevante Fragen wurden in der Literatur keine äquivalenten Deutungsmuster gefunden, so daß lediglich geprüft wird,

- H 1: ob die soziodemographischen Merkmale der Interviewer einen Effekt haben
(operational: signifikante Regressionskoeffizienten).

Der Einfluß dieses Typs von Interviewermerkmalen wird für in der Übersicht 1 (S.88ff., 2.1 und 2.2) angegebene abhängige Variablen überprüft.

- (2) spezifische Einstellungs- und Verhaltensweisen des Interviewers, wozu wir ebenfalls im Teil 2.1 exemplarisch einige Arbeiten genannt haben⁸.

Zur Frage nach der Richtung der Effekte dieser Merkmale wird in der Literatur eine positive Beziehung zwischen spezifischen Einstellungs- und Verhaltensweisen der Interviewer und denen der Befragten behauptet und (zum Teil) auch nachgewiesen.

Wollte man hinterfragen, warum es zu einer positiven Beziehung kommt, so wäre die Antwort in einer Anpassung des Befragten an die (von ihm vermuteten⁹) Einstellungs- und Verhaltensweisen des Interviewers zu suchen, aus der sich der Befragte positive Folgen (wie gegenseitige Bestätigung durch Ähnlichkeit, Attraktion usw.) erhofft.

Nicht klar ist allerdings, unter welchen Bedingungen Effekte von Interviewermerkmalen dieses Typs 2 ausgehen. Wir vermuten, daß spezifische

Übersicht 1: abhängige und unabhängige Variablen

unabhängige Variablen

abhängige Variablen

Frage	Kurzbezeichnung	Ausprägungen	1 Erziehungs- ziele	2 Oben-/Unten- Skala	3 subj. Schicht- einstufung	4 Zweitstimme	5 Parteisympa- thiescore	6 Postmateri- alismus-Sc.
(1) <u>s u b s t a n t i e l l e</u> UV's (Befragten-Charakteristika)								
(1.1) <u>sozio-demographische Merkmale</u>								
S6	Berufsprestige Befragter	Treiman-Score (18-78); wenn gegenwärtig kein Beruf, dann früherer Beruf (Prestigewert)	1	1	1	1	1	1
S5	Erwerbstätigkeit Befragter	1 = erwerbstätig 0 = nicht erwerbstätig	1	1	1	1	1	1
	Haushalts-Eink.	(Addition der Einzeleinkommen im HH)	1	1	1	1	1	
S39	Einkommen des Befragten	in DM	1	1	1	1	1	
S3	Schulbildung des Befragten	1 = kein Abschluß 2 = Volks/Hauptschule 3 = Mittlere Reife, Realschulabschluß 4 = Fachhochschulreife 5 = Abitur	1	1	1	1	1	1
S36	Alter des Befragten	in Jahren	1			1	1	1
I	Geschlecht des Befragten	1 = männlich 2 = weiblich	1			1	1	1

Frage	Kurzbezeichnung	Ausprägungen	1	2	3	4	5	6
22	Schicht-Selbst-einstufung des Befragten	1 = Unterschicht ... 5 = Oberschicht	1	---	---			
S46	oben/unten-Einst.	1 = oben ... 10 = unten	1	---	---			
<u>(1.2) Bezugsgruppen-Bindung</u>								
S45A	Gewerkschafts-mitgl. (DGB)	0 = nicht Mitglied 1 = Mitglied			1	1	1	
S37	Religionszugeh. des Befragten	1 = katholisch 0 = nicht katholisch	1			1	1	1
S38	Kirchgangshäufig.	1 = oft ... 5 = selten/nie	1			1	1	1
<u>(1.3) Kapital-Bindung</u>								
S20	Eigentum: Kapitalvermögen	1 = vorhanden 2 = nicht vorhanden		1	1			
2	Eigentum: Eigenheim	0 = keines (alte Codes 1-5,8) 1 = vorhanden (alte Codes 6-7)		1	1			
<u>(1.4) sonstige Befragtenmerkmale</u>								
32	Links-Rechts-Sk.	1 = links ... 10 = rechts				1	1	
34	Parteienskalometer	Differenz zwischen Sympathie für SPD minus der Sympathie für die CDU				1	---	
28	Post-Materialismus-Score des Befragten	Zusammenfassung der beiden wichtigsten Ränge zu 4 Typen: 1 = Materialisten, 2-3 = Mischtypen 4 = Post-Materialisten	1			1	1	---

Frage	Kurzbezeichnung	Ausprägungen	1	2	3	4	5	6
(2) <u>I n t e r v i e w e r - Merkmale als UV</u>								
(2.1) <u>soziodemographische Merkmale</u>								
S9	berufliche Stellung des Interviewers	1 = Arbeiter 2 = Beamte und Angestellte 3 = Akademiker, freie Berufe, Selbständige (außerhalb der Landwirtschaft)	3	3	3	3	3	3
S6	Schulbildung des Interviewers	siehe 1.1 (Schulbildung des Befragten)	3	3	3	3	3	3
S2	Alter des Int.	in Jahren	3	3	3	3	3	3
S1	Geschlecht des Interviewers	1 = männlich 2 = weiblich	3	3	3	3	3	3
(2.2) <u>Bezugsgruppen-Bindung</u>								
S10	Gewerkschaftsmitgl. (DGB)	0 = nicht Mitglied 1 = Mitglied		3	3	3	3	
S4	Religionszugeh. des Interviewers	1 = katholisch 0 = nicht katholisch	3			3	3	3
S5	Kirchgangshäufig.	siehe 1.2 (Kirchgangshäufigkeit d. Befragten)	3			3	3	3
(2.3) <u>allgemeine Interviewer-Merkmale</u>								
1	Zahl der 1979 durchgeführten Interviews	Anzahl	3	3	3	3	3	3

Frage	Kurzbezeichnung	Ausprägungen	1	2	3	4	5	6
S8	"zeitliche Beschränkung des Interviewers"	1 = zeitlich sehr beschränkt (alte Codes 1,7) 2 = zeitlich weniger beschr.(alte Codes 2,4,8) 3 = zeitlich frei verfügb.(alte Codes 3,5,6,9)	3	3	3	3	3	3
(2.4) <u>Einstellungs-/Verhaltensmerkmale der Interviewer</u>								
A	Erziehungsziele-index	Summenscore der im Text beschriebenen Items (analog zum Index Erz.ziele der Befragten)	3					
C	oben/unten Einst. des Interviewers	siehe 1.1 (Oben/Unten-Einstellung d. Befragten)		3	3			
D	Parteienskalo-meter	siehe 1.4 (Parteienskalometer d. Befragten)				3	3	
S12	Zweitstimme des Interviewers	2 = CDU und CSU 3 = SPD				3	3	
D	Symp. "Grüne"	+5 = sympathisch ... -5 = nicht sympathisch						3
(3) <u>Merkmale der Interview - Situation</u>								
II	"Dritte während des Interviews anwesend"	1 = nur Befragter und Interviewer anwesend 0 = Befragter, Interviewer und andere Personen anwesend	3	3	3	3	3	3

Einstellungs- und Verhaltensweisen des Interviewers dann von Bedeutung sind, wenn der Gegenstand, den sie betreffen, normativ verankert ist und der Befragte eine Vorstellung davon hat, wie relevante Antwortmuster aussehen und wie diese (mehrheitlich?) in der Bevölkerung verbreitet sind. (Dieses Kriterium ist zugegebenermaßen nicht exakt genug, um Fragen genau danach zu klassifizieren, ob nun ein Einfluß dieser Merkmale zu erwarten ist oder nicht.) Geprüft wird im weiteren,

H 2: ob zwischen den Interviewer-Erziehungszielen und den Befragten-Erziehungszielen, der Interviewer-Oben-Unten-Einstufung und der Befragten-Oben-Unten-Einstufung und Befragten-Subjektiven-Schichteinstufung, der Interviewer-Parteisympathie und Befragten-Parteisympathie und Befragten-Zweitstimme, der Sympathie des Interviewers für die "Grünen" und dem Befragten-Post-Materialismus¹⁰ eine positive Beziehung besteht.

(operational: signifikante, positive Regressionskoeffizienten).

Vgl. dazu auch wieder Übersicht 1 (S.88ff., 2.4).

(3) allgemeine (sonstige) Interviewermerkmale

wie zeitlich eingeschränkte Verfügbarkeit des Interviewers und Interviewererfahrung (Zahl der in einem bestimmten Zeitraum durchgeführten Interviews). Wie schon bei den Merkmalen des Typs 1 fehlt auch hier die intervenierende Theorie, die angibt, unter welchen Bedingungen und wie diese Merkmale wirken. Für abweichendes Verhalten wurde vermutet (und empirisch nachgewiesen: Schanz 1982), daß Befragte bei erfahrenen Interviewern mehr abweichendes Verhalten berichten (was den "wahren" Werten näherkommend interpretiert wurde). Die Interviewererfahrung dürfte wohl aber nicht generell über alle Fragetypen wirken. Wir prüfen daher,

H 3: ob allgemeine Interviewermerkmale einen Effekt auf die Angaben des Befragten haben

(operational: signifikante Regressionskoeffizienten).

Der Einfluß dieser Merkmale wird wieder über alle Modelle geprüft (vgl. Übersicht 1, S.88ff., 2.3).

(4) Interviewsituation

Hier soll der Einfluß der Anwesenheit/Abwesenheit Dritter während des Interviews geprüft werden. Wann derartige Effekte auftreten und in wel-

cher Richtung sie wirken, kann durch eine generelle intervenierende Theorie nicht spezifiziert werden. Es können auch keine Ableitungen für unsere spezifischen abhängigen Variablen gemacht werden. Effekte können vielleicht erwartet werden, wenn Dritte sanktions- oder kontrollfähig sind (vgl. hierzu theoretisch auch Granovetter 1978).

Wir prüfen,

- H 4: ob die Anwesenheit Dritter während des Interviews einen Einfluß hat
(operational: signifikanter Regressionskoeffizient).

5. Methodische Vorgehensweise

Wir gehen multivariat in vier Schritten vor:

- (1) Regressionsanalysen für alle abhängigen Variablen mit den jeweiligen substantiellen unabhängigen Variablen. Die Variablen, die in diesen ersten Schritt eingehen, sind in der Übersicht 1, S.88ff., mit einer "1" gekennzeichnet.
- (2) Regressionsanalysen für diejenigen substantiellen unabhängigen Variablen, die (im ersten Schritt) empirisch erklärungskräftig waren. (Als Kriterium gilt die Höhe des beta-Koeffizienten - .09 bzw. .05, siehe unten - und der Standardschätzfehler.)
- (3) Zu den Variablen von (2) werden in einem dritten Schritt die in Übersicht 1, S.88ff., mit einer "3" gekennzeichneten Interviewer-/Interviewmerkmale hinzugefügt¹¹.
- (4) In einem vierten Schritt wurden (wie von Schritt 1 zu Schritt 2) wenig erklärungskräftige Interviewer-/Interviewmerkmale aussortiert und die restlichen Variablen nochmals in eine Regressionsanalyse eingegeben.

Wir führen damit von Schritt 2 zu Schritt 3 bzw. 4 neue (Interviewer-) Variablen ein, deren Nichtberücksichtigung - technisch gesprochen - führen kann zu

- (a) verzerrten Parametern durch Spezifikationsfehler wegen fälschlicherweise weggelassener unabhängiger Variablen (vgl. Deegan 1976; Duncan 1975: 101-112; Rao und Miller 1971).

Hier muß davon ausgegangen werden, daß gilt: Je stärker die weggelassenen Variablen mit den im Modell enthaltenen Variablen korrelieren, desto mehr sind die Koeffizienten des falsch spezifizierten Modells verzerrt. Korrelieren die weggelassenen Variablen nicht mit den im Modell vorhandenen Variablen (was ein empirisch unwahrscheinlicher Fall ist), so sind die Koeffizienten unverzerrt;

- (b) einer Unterschätzung der erklärten Varianz der abhängigen Variablen. Korrelieren die weggelassenen Variablen nicht mit anderen unabhängigen Variablen, so wird die erklärte Varianz um so mehr unterschätzt, je stärker die Beziehungen der weggelassenen Variablen mit der abhängigen Variablen sind. Korrelieren die weggelassenen Variablen mit anderen unabhängigen Variablen des Modells, dann wird die erklärte Varianz des falsch spezifizierten Modells unterschätzt oder bleibt gleich, wenn die weggelassenen Variablen auch mit der abhängigen Variable korreliert sind. Die erklärte Varianz wird unterschätzt, wenn die weggelassenen Variablen einen Effekt auf die abhängige Variable haben.

Es wird daher im weiteren geprüft,

H 5: ob sich die Koeffizienten (beta-Werte) der substantiellen Modelle nach Hinzufügen der Interviewer-/Interviewmerkmale verändern; d.h. beispielsweise, daß die Beziehung zwischen den substantiellen unabhängigen Variablen und der abhängigen Variable nach Hinzufügen der Interviewer-/Interviewmerkmale gegen 0 geht;

H 6: ob die erklärte Varianz nach Hinzufügen der Interviewer-/Interviewmerkmale ansteigt.

6. Ergebnisse

6.1 Allgemeine Rahmendaten zu den Interviewern

Tabelle 1, S.96, zeigt zentrale demographische und Einstellungsmerkmale von Interviewern (die mindestens ein Interview realisieren) und Befragten im Vergleich. Unter den Interviewern sind eher Frauen, Jüngere und Ledige als bei den Befragten. Interviewer zeichnen sich durch eine gehobene Schulbildung und höhere berufliche Stellung aus - beides sind wohl Voraussetzung für den "Nebenberuf" Interviewer bei professionellen Instituten.

6.2 Multivariate Analysen

Die Ergebnisse zu den beschriebenen Regressionsanalysen sind in den Übersichten 2-7, S.98ff., zusammenfassend dargestellt. Die jeweils erste und zweite Spalte in den Übersichten 2-7 geben die Regressionskoeffizienten für die substantiellen unabhängigen Variablen (1. bzw. 2. Schritt) wieder. Die dritte Spalte enthält die Koeffizienten der ausgewählten substantiellen unabhängigen Variablen (des 2. Schrittes) und die der Interviewermerkmale. Spalte 4 enthält die ausgewählten Interviewermerkmale und die ausgewählten substantiellen unabhängigen Variablen.

Ein "+" in diesen Übersichten bedeutet, daß der jeweilige Koeffizient auf dem 5%-Niveau signifikant ist, ein "++" steht für eine Signifikanz auf dem 1%-Niveau. Dabei wurde für die substantiellen unabhängigen Variablen ein einseitiger Test, für die Interviewermerkmale ein zweiseitiger Test zugrunde gelegt, da im letztgenannten Falle die Beziehungen für die meisten Merkmale nicht a priori postuliert werden können. Wir gehen im folgenden auf jedes Modell (d.h. jede abhängige Variable) - auch im Hinblick auf die substantiellen unabhängigen Variablen, zu denen wir zum Teil recht spekulative Postulate präsentierten - kurz ein.

Liberaler Erziehungsziele (vgl. Übersicht 2)

Vier substantielle unabhängige Variablen (Alter, subjektive SchichtEinstufung, Haushaltseinkommen, Post-Materialismus-Score) erklären hier

Tabelle 1: Befragten- und Interviewermerkmale im Vergleich

	% Befragte (N=2955)	% Interviewer (N=430)
Geschlecht: Männer	46,0	38,1
Frauen	54,0	61,9
Alter: bis 32 Jahre	25,5	39,7
33-46 Jahre	27,2	29,3
47-60 Jahre	22,2	20,3
älter als 60 Jahre	25,2	10,4
Familienstand: verheiratet	62,7	65,5
ledig	17,8	25,5
Konfession: evangelisch	50,3	45,7
katholisch	42,1	42,6
keine	6,6	10,5
Schulabschluß: kein Abschluß	2,0	0,5
Volksschule	62,4	27,3
Mittlere Reife, Realschule	21,9	38,2
Fachhochschulreife	3,7	8,8
Abitur	10,0	25,4
Berufliche Stellung: Arbeiter	32,1	8,6
Beamte, Angestellte	53,3	84,5
Akademiker, freie Berufe	9,5	6,9
sonstige Selbständige		
Indexwert "liberale Erziehungsziele"		
bis 14 Punkte	0,6	0,6
15-21 Punkte	4,8	2,9
22-28 Punkte	26,6	21,4
mehr als 29 Punkte	68,1	74,9
Indexwert "Parteisympathie": pro CDU	39,6	35,1
ausgeglichen	11,6	13,3
pro SPD	48,8	51,6
Selbsteinstufung "Oben/Unten"		
1-2	2,9	0,9
3-4	15,2	6,9
5	19,0	9,1
6	33,9	35,3
7	17,2	25,7
8-10	11,7	22,0
Zweitstimme: SPD	(Basis: nur SPD- u.	67,6
CDU/CSU	CDU/CSU-Wähler)	32,4
	47,7	
	52,3	

Die %-Angaben addieren sich bisweilen nicht auf 100, weil etwa beim Familienstand nur ausgewählte Kategorien hier dargestellt wurden.

lediglich etwa 2% der Varianz der abhängigen Variablen. Der Konfessionsfaktor erweist sich als bedeutungslos; subjektive Schichteinstufung und Post-Materialismus hingegen wirken wie vermutet: Je höher die subjektive Schichteinstufung und der Post-Materialismus-Score, desto stärker vertreten die Befragten liberale Erziehungsziele.

Die Hinzunahme von Interviewermerkmalen zeigt, daß zwei Merkmale aus diesem Bereich erkennbar zu einer Varianzaufklärung beitragen: die berufliche Stellung des Interviewers und der Index "liberale Erziehungsziele" des Interviewers. Letztgenanntes (unsichtbares) Interviewermerkmal führt sogar zu einer Vervielfachung der erklärten Varianz. Der Effekt dieses Merkmals ist stärker als der aller anderen Variablen im Modell zusammengenommen. Die Tabelle 2 zeigt den Zusammenhang zwischen Interviewer-und Befragten-"Erziehungsziele" in bivariater Betrachtungsweise.

Tabelle 2: Index "liberale Erziehungsziele"

		B E F R A G T E R			N	
		schwach	mittel	stark		
I N T E R V I E W E R	schwach	510 53,5	193 20,3	250 26,2	953	$\bar{X} = 1,73$ SD = 0,85
	mittel	389 37,8	291 28,3	349 33,9	1029	$\bar{X} = 1,96$ SD = 0,85
	stark	227 25,6	210 23,7	448 50,6	885	$\bar{X} = 2,25$ SD = 0,84
		1126	694	1047	2867	

Gamma = .32 ; Tau b = .22 ; Zeilenprozent

Um den Effekt der beruflichen Stellung und der Schulbildung der Interviewer zu kontrollieren, haben wir diese Analysen der Erziehungsziele für Interviewer-Subgruppen (nach Schulbildung, beruflicher Stellung des Interviewers) nochmals gerechnet.

Übersicht 2: abhängige Variable - Index "liberale Erziehungsziele"

BEFRAGTEN - MERKMALE:

Alter	-.053	-.056	-.047	-.049 ⁺⁺
Geschlecht	.039			
Schulbildung	-.013			
Berufsprestige	-.012			
Einkommen	.031			
Haushaltseinkommen	-.055	-.040	-.029	-.023
Oben-/Unten-Skala	-.018			
Schichteistufung	.076	.064	.071	.071 ⁺⁺
Post-Materialismus-Score	.094	.087	.085	.082 ⁺⁺
Konfession	-.015			
Kirchgangshäufigkeit	-.025			

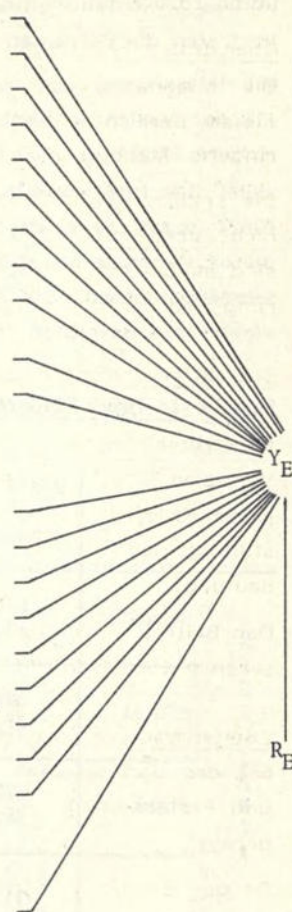
INTERVIEWER - MERKMALE:

Alter	.025	
Geschlecht	-.043	
Schulbildung	.003	
berufl.Stellung	-.082	-.086 ⁺⁺
Konfession	.017	
Kirchgangshäufigkeit	-.019	
Erziehungsziele	.279	.230 ⁺⁺
Anz. Interviews 1979	-.051	
Zeitverfügbarkeit	-.030	

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	-.039
--	-------

erklärte Varianz	.02	.02	.09	.10
------------------	-----	-----	-----	-----



Es zeigte sich, daß es für die einzelnen Subgruppen kaum erkennbare Abweichungen gibt; die erklärte Varianz und der jeweilige F-Wert sind stabil. Wir kommen auf die abhängige Variable "liberale Erziehungsziele" unten zurück.

Oben-Unten-Einstufung (vgl. Übersicht 3)

Mit "klassischen" Variablen wie Beruf und Einkommen kann hier immerhin 11% der Varianz erklärt werden. Alle anderen Merkmale haben weitaus geringere Effekte.

Die Erklärungskraft kann durch die Hinzunahme von Interviewermerkmalen nicht verbessert werden. Schwache Einflüsse von Interviewermerkmalen sind im Beitrag der Schulbildung des Interviewers, der Interviewererfahrung und der Oben-Unten-Einstufung des Interviewers angedeutet.

Subjektive Schichteinstufung (vgl. Übersicht 4)

Mit drei "klassischen" Merkmalen - Berufsprestige, Schulbildung und Haushaltseinkommen - kann hier zunächst 28% der Varianz der abhängigen Variablen erklärt werden. Diese Variablen sind in ihrem Gewicht nicht in jedem Falle mit den erklärungskräftigsten Variablen zur Oben-Unten-Einstufung identisch. Auch hier sind andere Merkmale nur von geringer Bedeutung.

Der Beitrag von Interviewermerkmalen ist gering, so daß auf den vorgesehenen vierten Schritt verzichtet wurde.

Zweitstimme (vgl. Übersicht 5)

68% der Varianz kann hier mit nur zwei Merkmalen - Links-Rechts-Skala und Parteisympathie - aufgeklärt werden. Alle anderen Koeffizienten sind gering.

Da der Beitrag der Interviewermerkmale gering ist, wurde auch hier auf den vierten Analyseschritt verzichtet. Relativ gesehen erklären von allen Interviewermerkmalen noch am meisten: Zweitstimme des Interviewers und Parteisympathie des Interviewers.

Übersicht 3: abhängige Variable - Oben-/Unten-Einstufung

BEFRAGTEN - MERKMALE:

Erwerbstätigkeit	.013			
Berufsprestige	.080	.122	.126	.129 ⁺⁺
Schulbildung	.062			
Einkommen	.112	.111	.105	.105 ⁺⁺
Haushaltseinkommen	.161	.181	.180	.179 ⁺⁺
Kapitalvermögen	-.011			
Hausbesitz	.048			
DGB-Mitgliedschaft	-.045			

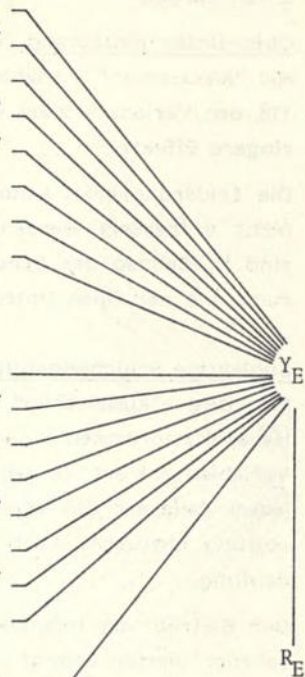
INTERVIEWER - MERKMALE:

berufl.Stellung	.029			
Schulbildung	-.064	-.057 ⁺⁺		
Alter	-.020			
Geschlecht	.001			
Oben-/Unten-Einstufung	.052	.053 ⁺⁺		
DGB-Mitgliedschaft	.022			
Anz.Interviews 1979	-.057	-.055 ⁺⁺		
Zeitverfügbarkeit	.029			

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	.014			
--	------	--	--	--

erklärte Varianz	.12	.11	.11	.12
------------------	-----	-----	-----	-----



Übersicht 4: abhängige Variable - subjektive Schichteinstufung

BEFRAGTEN - MERKMALE:

Erwerbstätigkeit	-.051
Berufsprestige	.217 .229 .220 ⁺⁺
Schulbildung	.230 .240 .234 ⁺⁺
Einkommen	.082
Haushaltseinkommen	.190 .222 .224 ⁺⁺
Kapitalvermögen	-.005
Hausbesitz	.026
DGB-Mitgliedschaft	-.091

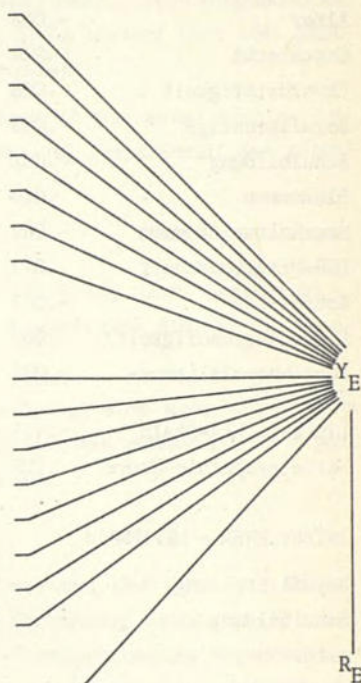
INTERVIEWER - MERKMALE:

berufl.Stellung	.042 ⁺
Schulbildung	.038 ⁺
Alter	.019
Geschlecht	.009
Oben-/Unten-Einstufung	-.016
DGB-Mitgliedschaft	.038 ⁺
Anz.Interviews 1979	-.051
Zeitverfügbarkeit	.063 ⁺⁺

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	.026
--	------

erklärte Varianz	.39 .28 .29
------------------	-------------



Übersicht 5: abhängige Variable - Zweitstimme (CDU/CSU - SPD)

BEFRAGTEN - MERKMALE:

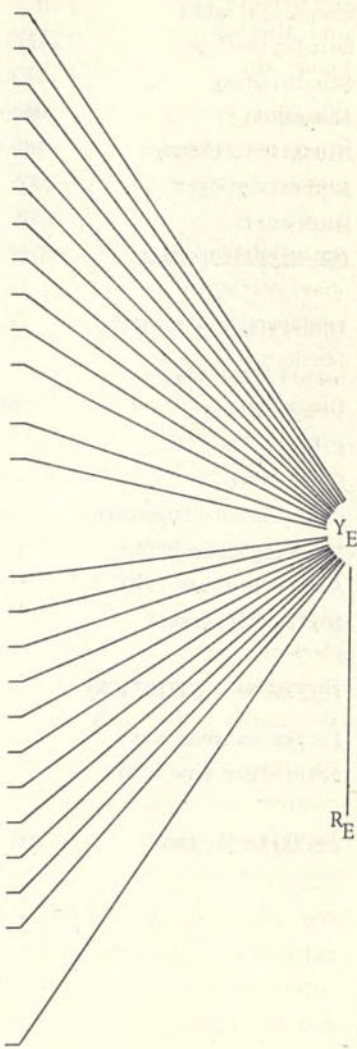
Alter	-.035		
Geschlecht	.024		
Erwerbstätigkeit	.005		
Berufsprestige	.035		
Schulbildung	-.032		
Einkommen	.026		
Haushaltseinkommen	-.062		
DGB-Mitgliedschaft	.071		
Konfession	-.072		
Kirchgangshäufigkeit	.069		
Post-Materialismus-Score	-.021		
Links-Rechts-Skala	-.135	-.138	-.137 ⁺⁺
Parteisympathie-Score	.705	.752	.752 ⁺⁺

INTERVIEWER - MERKMALE:

berufl. Stellung	.010		
Schulbildung	-.043 ⁺⁺		
Alter	-.035 ⁺⁺		
Geschlecht	.018		
Anz. Interviews 1979	-.017		
Zeitverfügbarkeit	.020		
DGB-Mitgliedschaft	.010		
Parteisympathie-Score	-.098 ⁺⁺		
Zweitstimme	.090 ⁺⁺		
Konfession	-.031 ⁺⁺		
Kirchgangshäufigkeit	-.023		

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	.010		
--	------	--	--



erklärte Varianz .71 .68 .69

Parteisympathie-Score (vgl. Übersicht 6)

Mit vier Variablen kann hier noch 29% der Varianz erklärt werden, wobei die erklärungskräftigsten Variablen der "Zweitstimme" mit denen der "Parteisympathie" auch nicht identisch sind (Geschlecht, DGB-Mitgliedschaft und Kirchengangshäufigkeit und Links-Rechts-Skala wirken sehr viel deutlicher auf die Parteisympathie als die Zweitstimme).

Die Hinzunahme von Interviewermerkmalen bewirkt nur einen leichten Anstieg der erklärten Varianz. Am deutlichsten ist der Beitrag der Interviewer-Parteisympathie.

Post-Materialismus (vgl. Übersicht 7)

Vier Merkmale - Alter, Berufsprestige, Schulbildung und Geschlecht - erklären hier 19% der Varianz. Der von uns vermutete Einfluß des Konfessionsfaktors ist gering.

Die Hinzunahme von Interviewermerkmalen erbringt eine kaum erkennbare Steigerung der erklärten Varianz. Allerdings ist der relativ starke Effekt des Alters des Interviewers hier hervorzuheben.

Interviewermerkmale im Überblick

Es konnte kein Einzelmerkmal identifiziert werden, das über alle abhängigen Variablen eine etwa gleiche relative Bedeutung hatte. Unsichtbare Merkmale, die mit der abhängigen Variablen (entsprechendes Merkmal beim Befragten) verbunden sind, erklären relativ gesehen mehr als etwa die sichtbaren Merkmale (wie Alter oder Geschlecht des Interviewers).

Da "berufliche Stellung" und "Interviewererfahrung" - die beide noch relativ viel erklären - ebenfalls für den Befragten "unsichtbar" sind, kann wohl davon ausgegangen werden, daß "unsichtbare Merkmale" das Befragtenverhalten noch am deutlichsten beeinflussen.

Die Wirkung von Interviewermerkmalen variiert mit der entsprechenden abhängigen Variablen. Es gibt offensichtlich keinen Standardset von Interviewermerkmalen, der quer über alle Variablen gleichgerichtet auf das Befragtenverhalten einwirkt.

Ein Vergleich der Koeffizienten der substantiellen unabhängigen Variablen ohne und mit Interviewermerkmalen zeigt, daß diese bemerkenswert stabil sind¹².

Übersicht 6: abhängige Variable - Parteisympathie-Score

BEFRAGTEN - MERKMALE:

Alter	-.062			
Geschlecht	.088	.082	.079	.080 ⁺⁺
Erwerbstätigkeit	.005			
Berufsprestige	-.035			
Schulbildung	-.029			
Einkommen	.052			
Haushaltseinkommen	-.076			
DGB-Mitgliedschaft	.093	.102	.102	.098 ⁺⁺
Konfession	-.072			
Kirchgangshäufigkeit	.163	.204	.206	.206 ⁺⁺
Post-Materialismus-Score	.071			
Links-Rechts-Skala	-.423	-.441	-.437	-.435 ⁺⁺

INTERVIEWER - MERKMALE:

berufl. Stellung	-.019			
Schulbildung	-.012			
Alter	-.051	-.040 ⁺		
Geschlecht	.007			
Anz. Interviews 1979	.002			
Zeitverfügbarkeit	.078			
DGB-Mitgliedschaft	-.043			
Parteisympathie-Score	.074	.080 ⁺⁺		
Zweitstimme	.007			
Konfession	-.032			
Kirchgangshäufigkeit	-.048	-.039 ⁺		

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	.003			
--	------	--	--	--

erklärte Varianz	.41	.29	.30	.30
------------------	-----	-----	-----	-----

Y_ER_E

Übersicht 7: abhängige Variable - Post-Materialismus-Score

BEFRAGTEN - MERKMALE:

Alter	-.28	-.27	-.26	-.26 ⁺⁺
Berufsprestige	.09	.08	.09	.09 ⁺⁺
Erwerbstätigkeit	-.03			
Schulbildung	.22	.23	.21	.22 ⁺⁺
Konfession	-.04			
Kirchgangshäufigkeit	.04			
Geschlecht	-.05	-.04	-.05	-.05 ⁺⁺

INTERVIEWER - MERKMALE:

berufl. Stellung	-.02			
Schulbildung	.03			
Alter	-.10	-.11 ⁺⁺		
Geschlecht	.02			
Konfession	-.04	-.04 ⁺		
Kirchgangshäufigkeit	-.04	-.04 ⁺⁺		
Anz. Interviews 1979	-.00			
Zeitverfügbarkeit	.00			
Parteisympathie "Grüne"	.05	.06 ⁺⁺		

INTERVIEW - SITUATION:

Dritte während des Interviews anwesend	.00
---	-----

erklärte Varianz	.19	.19	.20	.20
------------------	-----	-----	-----	-----

Y_ER_E

Abweichungen $> .01$ zeigen sich für die abhängige Variable "liberale Erziehungsziele" beim Haushaltseinkommen des Befragten (.03); für die abhängige Variable "Oben-Unten-Skala" beim Haushaltseinkommen (.02) und dem Berufsprestige des Befragten (.04); für die abhängige Variable "Schichteinstufung" beim Haushaltseinkommen (.03); für die abhängige Variable "Zweitstimme" bei der Parteisympathie des Befragten (.05); für die abhängige Variable "Parteisympathie" bei der Kirchengangshäufigkeit des Befragten (.04); und für die abhängige Variable "Post-Materialismus" beim Alter des Befragten (.02).

Ein Vergleich der erklärten Varianz (2. versus 4. Schritt) zeigt, daß durch Hinzufügen der Interviewermerkmale bei den Erziehungszielen als abhängige Variable 8% mehr an Varianz erklärt werden kann. Bei allen anderen abhängigen Variablen beträgt der Zuwachs an erklärter Varianz nicht mehr als 1%.

Beides zeigt, daß ein bedeutender Einfluß von Interviewer-/Interviewmerkmalen (außer bei der abhängigen Variable "liberale Erziehungsziele") auf das Befragtenverhalten nicht nachweisbar ist.

Exkurs: "Erziehungsziele"

Es stellt sich die Frage, warum besonders die liberalen Erziehungsziele des Interviewers solch einen starken Effekt auf die liberalen Erziehungsziele des Befragten haben. Hierfür scheinen uns zwei alternative Erklärungsmöglichkeiten in Betracht zu kommen.

Eine Erklärung könnte darin bestehen, daß die vorhandene Beziehung zwischen den beiden Variablen aufgrund eines Drittfaktors, und zwar der Tendenz, sozial wünschbar zu reagieren, existiert. Allerdings kann dieser Erklärungsvorschlag nicht empirisch überprüft werden, da zum einen soziale Wünschbarkeit nicht bei den Interviewern erhoben wurde, zum anderen sich in dieser Studie gezeigt hat, daß die verwendeten Items zur Messung sozialer Wünschbarkeit nicht sehr valide zu sein scheinen (vgl. Porst und Schmidt, 1983).

Eine zweite Erklärung könnte darin liegen, daß die Erziehungsziele, verglichen mit allen anderen ausgewählten Variablen, am stärksten sozial wünschbare Tatbestände, wie z.B. Wichtigkeit von Fleiß oder Kritikfähigkeit, ausdrücken. Damit dürfte die Valenz der Folgen des Antwort-

verhaltens und die Erwartung des Befragten, positive Belohnungen zu erhalten, stark steigen. Als Folgen eines multiplikativen Effekts von sozialer Wünschbarkeit und Valenz bzw. Erwartung würde ein Einfluß auf das Antwortverhalten zu erwarten sein (vgl. hierzu Abbildung 3, S.108).

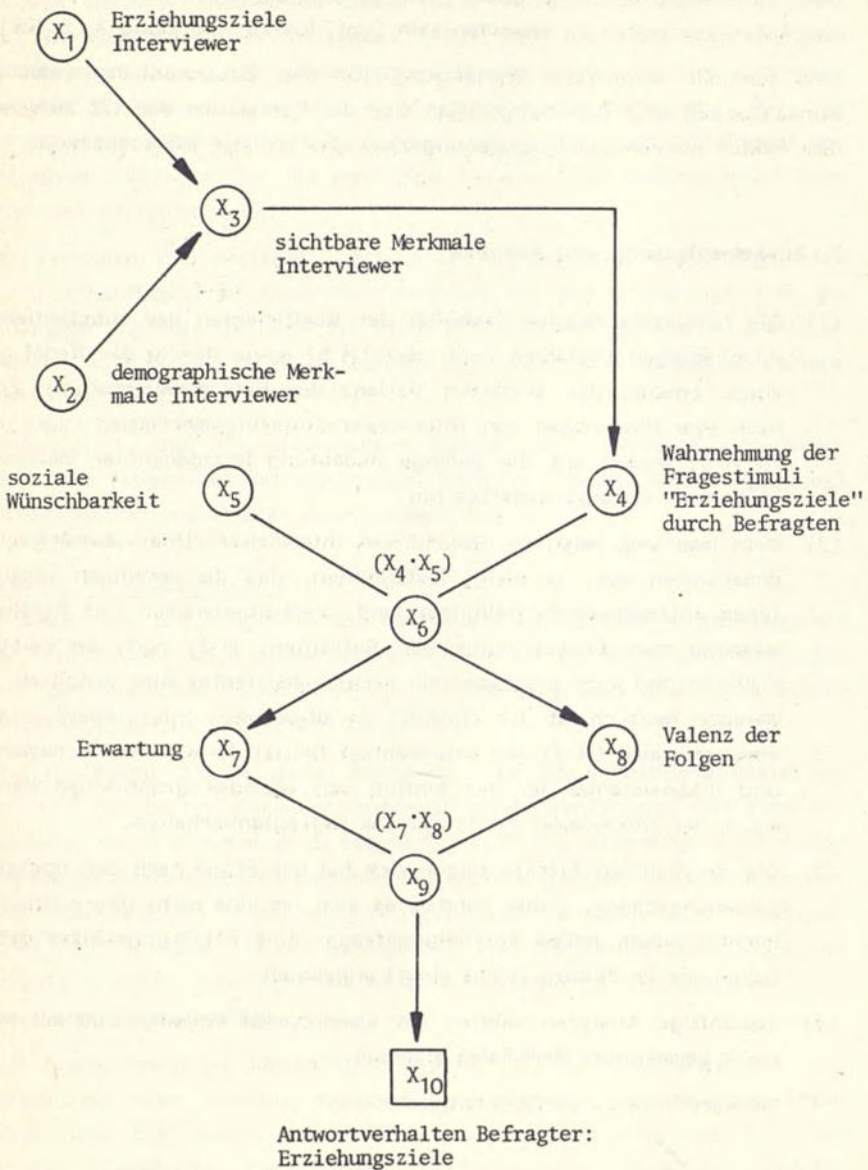
Eine partielle empirische Bestätigung für den Zusammenhang zwischen Wünschbarkeit und Erziehungszielen war die Korrelation von .26 zwischen dem Faktor konventionelle Erziehungsziele und sozialer Wünschbarkeit.

7. Zusammenfassung und Ausblick

- (1) Die berichtete relative Stabilität der Koeffizienten der substantiellen unabhängigen Variablen (vgl. dazu H 5) sowie der in der Regel geringe Anstieg der erklärten Varianz des Befragtenverhaltens auch nach dem Hinzufügen von Interviewer-/Interviewmerkmalen (vgl. dazu H 6) weisen auf die geringe Bedeutung letztgenannter Merkmale im Rahmen unseres Ansatzes hin.
- (2) Geht man vom relativen Gewicht von Interviewer-/Interviewmerkmalsdimensionen aus, so bleibt festzuhalten, daß die jeweiligen spezifischen Interviewer-Einstellungen und -Verhaltensweisen (als Parallelmessung zum Fragestimulus der Befragten; H 2) noch am meisten erklären und auch insgesamt ein recht konsistentes Bild vermitteln. Weniger deutlich ist der Einfluß von allgemeinen Interviewer-/Interviewmerkmalen (H 3) und Anwesenheit Dritter (H 4); noch schwächer und inkonsistenter ist der Einfluß von soziodemographischen Merkmalen der Interviewer (H 1) auf das Befragtenverhalten.
- (3) Die deutlichsten Effekte zeigen sich bei der Frage nach den liberalen Erziehungszielen. Dabei handelt es sich um eine nicht überprüfbare, leichte, wenig heikle Einstellungsfrage. Eine Erklärungsskizze dafür haben wir im Exkurs (siehe oben) angeboten.
- (4) Zukünftige Analysen müßten mit komplexeren Modellen und mit hier nicht gemessenen Merkmalen arbeiten.

So wäre etwa zu berücksichtigen:

Abbildung 3: Modell 2 - erweitertes Modell



- (a) die Interaktionen zwischen den unabhängigen Variablen (Befragtenmerkmale und Interviewermerkmale) addieren sich möglicherweise nicht in ihrer Wirkung auf das Antwortverhalten;
- (b) das Problem multipler Indikatoren (und damit die Kontrolle zufälliger und systematischer Meßfehler)¹³;
- (c) die Mehrzahl der hier angewandten Variablen weist im strengen Sinn kein Intervallskalenniveau auf; es sollten daher alternative Verfahren mit geringeren Skalenvoraussetzungen eingesetzt werden, um die Stabilität und Robustheit unserer Ergebnisse zu überprüfen.
- (d) stärkerer Bezug zu einer allgemeinen Theorie (indem die dort spezifizierten Variablen empirisch erfaßt werden) (vgl. das komplexe Modell bei Esser 1983, in diesem Band).

Anmerkungen

- 1 Im wesentlichen wird zur "Interviewsituation" erfaßt, ob der Befragte während des Interviews mit dem Interviewer allein ist oder ob noch Dritte anwesend sind. Wir werden diesen Einfluß - siehe unten - hier auch überprüfen.
- 2 Es sei noch auf Arbeiten zu verbaler/nicht-verbaler Verstärkung im Rahmen der Forschungen zum Interview verwiesen: Marquis 1970; Phillips und Clancy 1972; Erbslöh 1973; Cannell, Oksenberg und Converse 1979.
- 3 Zwischen heikel und schwierig kann bisweilen nur aufgrund von Plausibilität differenziert werden, da für beide Fragetypen dieselben Indikatoren gelten (Koolwijk 1969: Zahl der Verweigerungen).
- 4 Die dem Summenschore zugrundeliegende Annahme gleicher Gewichte ist insofern gerechtfertigt, als sich in einer Faktorenanalyse die Faktorladungen der Items auf dem entsprechenden Faktor in etwa als gleich erwiesen.
- 5 Hier kann die Oben-Unten-Skala der Interviewer als Parallelmessung akzeptiert werden; "subjektive Schichteinstufung" wurde bei den Interviewern nicht erhoben.
- 6 Insofern die Richtung der Effekte dieser Merkmale nicht angegeben wird, ist diese der zitierten Literatur zu entnehmen. Unser Anliegen hier ist nicht die Konkretisierung und Prüfung inhaltlicher Modelle.
- 7 Demographische Merkmale der Interviewer wurden - mit Ausnahme des "Einkommens" - relativ vollständig und mit der Erhebung dieser Merkmale bei den Befragten identisch erfaßt.

- 8 Aus Zeit- und Kostengründen wurde im ALLBUS nur eine beschränkte Ad-hoc-Auswahl von Indikatoren abgefragt.
- 9 Üblicherweise wird angenommen, daß Befragte die relevante Einstellungen- und Verhaltensweisen der Interviewer erkennen aufgrund derer sichtbarer Merkmale und den verbalen und nicht-verbalen Stimuli, die der Interviewer übermittelt. Zu letzterem gehört auch die Ausfüllung des vorhandenen Interpretationsspielraums des Interviewers etwa bei Präzisierungsversuchen von Fragen, bei Rückfragen und bei Protokollierung unklarer Antworten.
- 10 Dies soll hier als "related measure" verstanden werden.
- 11 Für einige abhängige Variablen war hier bereits ein Endpunkt erreicht, da die hinzugefügten Merkmale dieses Schrittes in einigen Fällen nur sehr geringe beta-Koeffizienten erreichten.
- 12 Auf eine Tabelle hierzu wird verzichtet; vgl. dazu die Übersichten 2-7, S.98ff., jeweils erste versus letzte Spalte für die "Befragten-Merkmale".
- 13 Zum Problem der zufälligen und systematischen Kontrolle von Meßfehlern sei verwiesen auf Dieter Hermann 1982: Kontextuelle Interviewereffekte. Gezeigt anhand der Daten des ALLBUS 1980.

Literatur

- Anger, Hans, 1969: Befragung und Erhebung. S.567-617 in: Carl F. Graumann (Hrsg.), Handbuch der Psychologie 7/1 (Sozialpsychologie). Göttingen: Hogrefe.
- Atteslander, Peter und Hans-Ulrich Kneubühler, 1975: Verzerrungen im Interview. Zu einer Fehlertheorie der Befragung. Opladen: Westdeutscher Verlag.
- Bargel, Tino, 1979: Überlegungen und Materialien zu Wertdisparitäten und Wertwandel in der BRD. S.147-184 in: Helmut Klages und Peter Kmiecik (Hrsg.), Wertwandel und gesellschaftlicher Wandel. Frankfurt, New York: Campus.
- Benney, Mark, David Riesman und Shirley A. Star, 1956: Age and Sex in the Interview. American Journal of Sociology, 62: 143-152.
- Berekhoven, Ludwig, Karl G. Specht, Keith Walthelm und F. Wimmer 1975: Zur Genauigkeit mündlicher Befragungen in der Sozialforschung. Bern: Lang.
- Berger, Manfred, 1977: Stabilität und Intensität von Parteineigung. Politische Vierteljahresschrift, 18: 501-509.
- Blankenship, Albert, 1940: The Effect of the Interviewer upon the Response in Public Opinion Poll. Journal of Consulting Psychology, 4: 134-136.
- Bradburn, Norman M., Seymour Sudman, Edward Blair und Carol Stocking, 1978: Question Threat and Response Bias. Public Opinion Quarterly, 42: 221-234.

- Cannell, Charles F., Lois Oksenberg und Jean M. Converse, 1979: Experiments in Interviewing Techniques. Field Experiments in Health Reporting 1971-1977. Ann Arbor, Michigan: Institute for Social Research.
- Cosper, Ronald, 1969: Interviewer Bias in a Study of Drinking Practices. *Quarterly Journal of Studies on Alcohol*, 30: 152-157.
- Deegan, John W., 1976: The Consequences of Model Misspecification in Regression Analysis. *Multivariate Behavioral Research*, 11: 237-248.
- Dijkstra, Will, o.J.: Response Bias in the Survey Interview. An Approach from Balance Theory. Amsterdam: Free University.
- Dijkstra, Will und E. Elsinga, 1979: Some Determinants of Interviewing Bias. Amsterdam: Free University (unveröff. Manuskript).
- Dohrenwend, Barbara S., 1969: Interviewer Biasing Effects: Toward a Reconciliation of Findings. *Public Opinion Quarterly*, 33: 121-125.
- Dohrenwend, Barbara S., John Colombotes und Bruce Dohrenwend, 1968: Social Distance and Interviewer Effects. *Public Opinion Quarterly*, 32: 410-422.
- Duncan, Otis D., 1975: Introduction to Structural Equation Models. New York: Academic Press.
- Erbslöh, Eberhard, 1973: Die unkontrollierte Einflußnahme des Interviewers im Forschungsinterview - dargestellt am Beispiel der Ergebnisse eines Tests zum Messen verbaler Intelligenz. S.1-67 in: René König und Erwin K. Scheuch (Hrsg.), *Kölner Beiträge zur Sozialforschung und angewandten Soziologie*. Meisenheim: Hain.
- Esser, Hartmut, 1975: Soziale Regelmäßigkeiten des Befragtenverhaltens. Meisenheim: Hain.
- Esser, Hartmut, 1983: Determinanten des Interviewer- und Befragtenverhaltens: Probleme der theoretischen Erklärung und empirischen Untersuchung von Interviewereffekten (in diesem Band).
- Falter, Jürgen W., 1977: Theoretische, methodologische und empirische Probleme einer Validierung des Konstrukts 'Parteiidentifikation' für die Bundesrepublik Deutschland. *Politische Vierteljahresschrift*, 18: 476-500.
- Feldman, Jakob J., Herbert H. Hyman und C.W. Hart, 1951: A Field Study of Interviewer Effects on the Quality of Survey Data. *Public Opinion Quarterly*, 15: 734-761.
- Ferber, Robert und Hugh G. Wales, 1952: Detection and Correction of Interviewer Bias. *Public Opinion Quarterly*, 16: 107-127.
- Freeman, John und Edgar W. Butler, 1976. Some Sources of Interviewer Variance in Surveys. *Public Opinion Quarterly*, 40: 79-91.
- Freitag, Carl B. und John R. Barry, 1974: Interaction and Interviewer Bias in a Survey of the Aged. *Psychological Reports*, 34: 771-774.
- Granovetter, Marc, 1978: Threshold Models of Collective Behavior. *American Journal of Sociology*, 83: 1420-1443.
- Hirschi, Travis und Hanan Selvin, 1976: *Delinquency Research*. New York, London: Free Press.

- Hodge, Robert W. und Donald J. Treiman, 1968: Class Identification in the United States. *American Journal of Sociology*, 73: 535-547.
- Hummell, Hans J. und Rolf Ziegler, 1976: Korrelation und Kausalität. Band 1. Stuttgart: Enke.
- Hyman, Herbert H., 1954: *Interviewing in Social Research*. Chicago: University Press.
- Inglehart, Ronald, 1979: Wertwandel in den westlichen Gesellschaften: Politische Konsequenzen von materialistischen und postmaterialistischen Prioritäten. S.279-316 in: Helmut Klages und Peter Kmiecik (Hrsg.), *Wertwandel und gesellschaftlicher Wandel*. Frankfurt, New York: Campus.
- Jackman, Mary R. und Robert W. Jackman, 1973: An Interpretation of the Relation between Objective and Subjective Social Status. *American Sociological Review*, 38: 569-582.
- Jöreskog, Karl G. und Dag Sörbom, 1981: The Use of LISREL in Sociological Model Building. S.179-200 in: Edgar F. Borgatta und Douglas N. Jackson (Hrsg.), *Factor Analysis and Measurement in Sociological Research*. London: Sage.
- Johnson, Weldon T. und John D. Delamater, 1976: Response Effects in Sex Surveys. *Public Opinion Quarterly*, 40: 165-181.
- Kaltefleiter, Werner und Peter Nißen, 1980: *Empirische Wahlforschung. Eine Einführung in Theorie und Technik*. Paderborn, München, Wien, Zürich: Schöningh.
- Kaufmann-Mall, Klaus, 1977: *Kognitiv-hedonistische Theorie menschlichen Verhaltens*. Bern: Huber.
- Klingemann, Hans D., 1978: Interviewereffekte bei der Protokollierung von Antworten auf offene Fragen. Materialien zum Vortrag auf der Herbsttagung der Sektion 'Methoden' der Deutschen Gesellschaft für Soziologie. Mannheim (unveröff. Manuskript).
- Klingemann, Hans D. und Charles L. Taylor, 1977: Einstellungskomponenten der Wahlentscheidung bei Bundestagswahlen in Deutschland. *Politische Vierteljahresschrift*, 18: 301-347.
- Kluegel, James R., Royce Singleton Jr. und Charles E. Stasner, 1977: Subjective Class Identification: A Multiple Indicator Approach. *American Sociological Review*, 42: 599-611.
- Koolwijk, Jürgen van, 1969: Unangenehme Fragen. Paradigma für die Reaktionen des Befragten im Interview. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 21: 864-875.
- Küchler, Manfred, 1977: Was leistet die empirische Wahlsoziologie? Eine Bestandsaufnahme. *Politische Vierteljahresschrift*, 18: 145-168.
- Küchler, Manfred, 1982: Eine sozio-demographische Beschreibung der Träger postmaterialistischer Einstellungen (in diesem Band).
- Lakatos, Imre, 1970: Falsification and the Methodology of Scientific Research Programs. S.91-196 in: Imre Lakatos und Alan Musgrave (Hrsg.), *Criticism and the Growth of Knowledge*. Cambridge: University Press.

- Lutynska, Krystyna, 1969: Third Persons in Sociological Interviews and Their Influence on the Respondents' Replies. *The Polish Sociological Bulletin*, 2: 139-145.
- Marquis, Kent H., 1970: Effects of Social Reinforcement on Health Reporting in the Household Interview. *Sociometry*, 33: 203-215.
- Pappi, Franz U., 1977: Sozialstruktur, gesellschaftliche Wertorientierungen und Wahlabsicht. *Politische Vierteljahresschrift*, 18: 195-229.
- Pfeil, Elisabeth und Jürgen Friedrichs, 1965: Überlegungen zum Problem rollenbedingter Interviewantworten. *Zeitschrift für Markt- und Meinungsforschung*, 8: 1877-1884.
- Phillips, Derek L. und Kelvin J. Clancy, 1972: 'Modelling Effects' in Survey Research. *Public Opinion Quarterly*, 36: 246-253.
- Porst, Rolf und Peter Schmidt, 1983: Analyse ausgewählter Meßinstrumente des ALLBUS 1980 (Arbeitstitel).
- Radtke, Günter D., 1977: Zur Prognosekompetenz der empirischen Wahlforschung. *Politische Vierteljahresschrift*, 18: 669-688.
- Rao, Potluri und Roger L. Miller, 1971: *Applied Econometrics*. Belmont, Calif.: Wadsworth Publishing.
- Sandberger, Johann U. und Tino Bargel, 1980: Überlegungen und Materialien zu subjektiven Indikatoren im internationalen Vergleich. S.155-212 in: Hans-Joachim Hoffmann-Novotny (Hrsg.), *Soziale Indikatoren im internationalen Vergleich*. Soziale Indikatoren VIII. Frankfurt, New York: Campus.
- Schanz, Volker, 1982: Interviewereffekte im Rahmen einer Befragung zu Trink- und Fahrgewohnheiten. Erscheint in: Heiner Meulemann und Karl-Heinz Reuband (Hrsg.), *Zur Rekonstruktion sozialer Realität bei der Erhebung und Analyse von Interviews* (Arbeitstitel).
- Smith, Harry L. und Herbert H. Hyman, 1950: The Biasing Effect of Interviewer Expectations on Survey Results. *Public Opinion Quarterly*, 14: 491-506.
- Stember, Ch. Herbert und Herbert H. Hyman, 1949: Interviewer Effects in the Classification of Responses. *Public Opinion Quarterly*, 13: 669-682.
- Sudman, Seymour und Norman M. Bradburn, 1974: *Response Effects in Surveys. A Review and Synthesis*. Chicago: Aldine.
- Sudman, Seymour, Norman M. Bradburn, Edward Blair und Carol Stocking, 1979: Role of Interviewer Expectations. S.51-63 in: Norman M. Bradburn und Seymour Sudman (Hrsg.), *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass.
- Taietz, Philip, 1962: Conflicting Group Norms and the 'Third' Person in the Interview. *American Journal of Sociology*, 68: 97-104.
- Weiss, Carol H., 1975: Interviewing in Evaluation Research. S.355-395 in: E.L. Struening und M. Guttentag (Hrsg.), *Handbook of Evaluation Research*. Beverly Hills. Sage.
- Williams, Allan J. Jr., 1964: Interviewer Respondent Interaction: A Study of Bias in the Information Interview. *Sociometry*, 27: 338-352.

ALLBUS 1980: STICHPROBENPLAN UND GEWICHTUNG

Hans-Peter Kirschner

1. Einleitung

Der ALLBUS 1980 gehört in bezug auf die Datenerhebungstechniken zu einem Typ bundesweiter Umfragen, der etwa seit 1973 von der Mehrzahl der kommerziellen Markt- und Meinungsforschungsinstitute in der Bundesrepublik fast ausschließlich und in zunehmender Weise routinisiert eingesetzt wird, wenn mit Hilfe einer Zufallsstichprobe ein repräsentativ angelegter Querschnitt der Bevölkerung erstellt werden soll. Die Auftraggeber sind dabei in aller Regel nicht-akademischen Institutionen zuzurechnen. Das Auftragsvolumen seitens der akademisch verfaßten Sozialwissenschaften ist demzufolge vergleichsweise sehr klein, und dies gilt, obwohl alle größeren Befragungen aus diesem Bereich mangels eigener Feldorganisation nur mit Hilfe privatwirtschaftlich verfaßter Meinungsforschungsinstitute durchgeführt werden können.

Eine solche Situation bestimmt selbstverständlich die Art der Dokumentation, insbesondere die des Stichprobenplans, der allen Untersuchungen dieses Typs zugrunde liegt: Die entsprechenden Veröffentlichungen sind eindeutig zugeschnitten auf die Mehrheit der Auftraggeber, so daß zwangsläufig Informationsbedürfnisse auf seiten der Sozialwissenschaften bisher unbefriedigt blieben.

Die vorliegende Arbeit soll ein Versuch sein, diese Informationslücke so weit wie möglich zu füllen. Sie wendet sich in erster Linie an Sozialwissenschaftler, die an den Methoden der Datenerhebung und ihren Konsequenzen interessiert sind. Es werden keine tiefergehenden mathematischen Vorkenntnisse vorausgesetzt und deshalb wird durchweg auf Beweise auch der bisher nicht bekannten stichprobentheoretischen Resultate verzichtet.

Wie eingangs gesagt, ist der ALLBUS 1980 - zumindest was den Stichprobenplan betrifft - Repräsentant für eine ganze Klasse von bundesweit angelegten Mehrthemenumfragen. Die meisten der folgenden Resultate zum Stichprobendesign und zur Gewichtung sind daher so allgemein formuliert,

daß sie ihre Gültigkeit auch für andere Studien gleichen Typs behalten. Im Vordergrund stehen dabei

- eine "klassische" Fehlerrechnung zu Mittelwertschätzungen im Rahmen eines mathematischen Ziehungsmodells (Abschnitte 4 und 5) und
- die Systematisierung von Gewichtungseffekten - unter Einschluß von Aussagen zur Schätzgenauigkeit - zweier in der Praxis häufig verwendeter Gewichtungsarten (Abschnitt 9).

Im einzelnen enthalten die Abschnitte 2 und 3 eine kommentierende Beschreibung des Stichprobenplans und seiner Umsetzung in die Praxis. Die Darstellung beschränkt sich bewußt auf die zentralen Eigenschaften dieses Plans und die entscheidenden Schritte zur praktischen Durchführung. So wird z.B. darauf verzichtet, Feinheiten im Aufbau diverser Datensätze zu diskutieren.

In den beiden formaler gehaltenen Abschnitten 4 und 5 wird innerhalb eines idealisierenden mathematischen Modells für den Stichprobenplan ein Ausdruck für die mittlere quadratische Abweichung eines auch für die Praxis grundlegenden Schätzverfahrens für Mittelwerte angegeben; zur Erläuterung der allgemeinen Ergebnisse wird für die Kohorte der 30- bis 39jährigen Männer in der Grundgesamtheit des ALLBUS 1980 anschließend angegeben, wie groß der - theoretische - Schätzfehler maximal sein kann.

Gegenstand der abschließenden drei Abschnitte ist es, die systematisierende Darstellung von Gewichtungseffekten zu verknüpfen mit der Diskussion über die erreichbare Genauigkeit zweier Schätzverfahren, die in der Praxis auf von nicht-stichprobenneutralen Ausfällen betroffene Stichproben häufig angewendet werden. Es ergeben sich dabei Kriterien, die als Entscheidungshilfe herangezogen werden können, wenn sich die Frage stellt, ob eine Auswertung mit den vorgeschlagenen Gewichten gerechnet werden soll oder nicht.

In diesem Zusammenhang sei darauf hingewiesen, daß in dieser Arbeit Gewichtungseffekte nur für die Schätzung von Anteilswerten behandelt werden. Damit ist zwar ein wesentlicher Teil multivariater Analysetechniken, nämlich solche, die sich ausschließlich auf relative Häufigkeiten stützen, abgedeckt; jedoch bleiben Fragen für diejenigen Verfahren offen, die mit den Werten der Variablen direkt arbeiten, wie z.B. Regressions- oder Varianzanalysen. Eine eingehendere Diskussion mußte in dem vorge-

gebenen Rahmen unterbleiben, da sie zwingend mit tiefergehenden mathematischen Ideen zur Modellbildung verknüpft ist. Im übrigen sind viele Problemstellungen dieses Themenkreises noch als ungelöst zu betrachten.

Die Durchnumerierung der Abschnitte ist fortlaufend; dabei sind die Abschnitte mit einer Dezimalpunkt-Numerierung nicht als Unterabschnitte zu verstehen. Diese Art der Bezeichnung wurde allein der besseren Lesbarkeit wegen eingeführt. Die Numerierung der Tabellen (Figuren) ist innerhalb der Abschnitte ebenfalls fortlaufend; Tab.8.3 ist also z.B. die dritte Tabelle in Abschnitt 8. Demselben Schema folgt die Kennzeichnung der Formeln, nur daß diese stets in Klammern steht; so hat die erste Formel in Abschnitt 4 die Nummer (4.1).

2. Die Grundgesamtheit

Von seiner Zielsetzung her sollte der ALLBUS 1980 ganz allgemein die Befragung eines repräsentativen Querschnitts der Bevölkerung der Bundesrepublik Deutschland und West-Berlins zu Wertorientierungen und Einstellungen, Meinungen und Verhalten darstellen (vgl. ZUMA 1979: 1f.). Vor Beginn der Untersuchung war also sowohl der Begriff "Bevölkerung", d.h. "umfassende Population", praktikabel zu definieren, als auch über den Stichprobenplan die Repräsentativität des Querschnitts sicherzustellen.

In der amtlichen Statistik wird durch den Begriff "Wohnbevölkerung" der allergrößte Teil der in der Bundesrepublik und West-Berlin lebenden Population erfaßt. Nicht abgedeckt wird lediglich die Gruppe der Personen - incl. Familienangehörigen -, die ausländischen Stationierungsstreitkräften sowie ausländischen diplomatischen und konsularischen Vertretungen angehören (vgl. Statistisches Bundesamt 1979:15). Zur Wohnbevölkerung gehören dagegen

- nahezu alle in der Bundesrepublik und West-Berlin lebenden Ausländer (also alle Personen, die nicht Deutsche im Sinne des Art.116, Abs.1 GG, sind) und
- alle Personen, die im Bereich einer Anstalt wohnen und keinen eigenen Haushalt führen ("Anstaltsbevölkerung").

In der Wohnbevölkerung sind also sicherlich alle Individuen enthalten, an die sich die Befragung ALLBUS 1980 plausiblerweise überhaupt wenden konnte, wobei allerdings

- Personen, die der deutschen Sprache nicht mächtig waren, kaum mit Hilfe des vorliegenden Fragebogens hätten befragt werden können und
- Personen aus der Anstaltsbevölkerung praktisch nur sehr schwer erreichbar sind.

Konsequenterweise kamen als potentiell zu Befragende nur solche Personen mit deutscher Staatsangehörigkeit in Betracht, die zum Zeitpunkt der Befragung in Privathaushalten¹ lebten. Des weiteren wurde, einer bei repräsentativen Bevölkerungsumfragen weit verbreiteten Übung folgend, dieser Kreis nochmals auf mindestens 18 Jahre alte Personen beschränkt.

Die der Untersuchung zugrundeliegende Grundgesamtheit wurde dementsprechend wie folgt definiert:

Alle Personen mit deutscher Staatsangehörigkeit, die zum Zeitpunkt der Befragung in der Bundesrepublik und in West-Berlin in Privathaushalten leben² und die spätestens am 1.1.1962 geboren wurden.

Diese Festlegung beschreibt offenbar keine feste Zielpopulation: Strenggenommen verändert sich die Menge derjenigen Menschen, auf welche die definierenden Kriterien für die Grundgesamtheit zutreffen, von Zeitpunkt zu Zeitpunkt, wie nahe diese auch beieinander liegen, insbesondere also von Tag zu Tag. Insgesamt kann eine solche Fluktuation jedoch vernachlässigt werden, da der Durchschnitt z.B. aller "Tages-Grundgesamtheiten", gebildet über alle 54 Tage der Feldzeit (7.1.80 - 29.2.80), sich nur unwesentlich von den einzelnen "Tages-Grundgesamtheiten" unterscheidet.

Der ALLBUS 1980 hat demnach in hohem Maße den erwünschten Charakter einer "Momentaufnahme".

Im Gegensatz zur Vorgabe "umfassende Population", die sich durch Abgrenzungen mit Hilfe von Begriffen der amtlichen Statistik recht einfach und auf sinnvolle Weise im Hinblick auf ihre Operationalität präzisieren läßt, ist die Vorgabe "Repräsentativität des Querschnitts" von weitaus höherer Komplexität. Dies darzustellen ist die Aufgabe insbesondere des Abschnitts 4.

3. Der Stichprobenplan

3.1 Alternativen zum Stichprobenplan des ALLBUS 1980

Zur Bereitstellung eines repräsentativen Querschnitts von Befragungspersonen aus der Grundgesamtheit waren drei verschiedene Methoden in der Diskussion, und zwar

- die Erstellung einer repräsentativen Quotenstichprobe,
- die Erstellung einer Zufallsstichprobe über die Einwohnermeldedateien der Gemeinden oder
- die Erstellung einer Zufallsstichprobe unter Verwendung des von der Mehrzahl der kommerziellen Markt- und Meinungsforschungsinstitute in der Bundesrepublik verwendeten mehrstufigen Stichprobenplans.

Der ALLBUS 1980 sollte seiner Zielsetzung nach u.a. "für die sozialwissenschaftliche Umfrageforschung in der Bundesrepublik überhaupt Modellcharakter haben" (vgl. ZUMA 1979: 3). Dies mußte sich selbstverständlich auch auf die Erhebungsmethode zu dieser von der Fallzahl her großen und von der Thematik her vielgestaltigen Umfrage beziehen. Dementsprechend fiel eine Quotenauswahl von vornherein aus, da ein solches Vorgehen typischerweise bei Mehrthemenbefragungen Schwächen allein schon deswegen aufweist, weil es praktisch unmöglich ist, aktuelle Quotierungsdaten zu solchen Quotierungsmerkmalen zu bekommen, die zu möglichst allen Untersuchungsmerkmalen eine hohe Korrelation besitzen. Und gerade dies wäre bei den Genauigkeitsansprüchen an diese Studie zu fordern gewesen (vgl. auch Neubäumer 1982: 181ff.). Man kann natürlich einwenden, daß auch bei Zufallsstichproben Angaben über die Genauigkeit der Schätzungen von Populationscharakteristika wegen der hohen Ausfallrate, die auch hier zu erwarten war, nicht sehr vertrauenswürdig sind. Dies trifft auch sicherlich zu, wenn die Gesamtheit der schließlich realisierten Interviews wegen systematischer Verzerrungen keinesfalls als "Stichprobe aus der Stichprobe" angesehen werden kann, eine Argumentation also innerhalb eines klassischen probabilistischen Ziehungsmodells, das in aller Regel die uneingeschränkte Verfügbarkeit der Befragungspersonen voraussetzt, mit vielen Unsicherheiten behaftet ist und somit Genauigkeitsangaben z.B. mit Hilfe von Varianzen zweifelhaft werden (vgl. Abschnitt 7). Jedoch ist zu bedenken, daß diese extreme Situation selbst

für eine große Stichprobe keineswegs immer eintreten muß und daß insbesondere für Teilpopulationen oft damit gerechnet werden kann, eine durch die Ausfälle relativ wenig "beschädigte" (Teil-)Stichprobe zu erhalten. Die Ableitung von Fehlerabschätzungen für bestimmte Arten von Schätzungen aus einem mathematischen Ziehungsmodell ist dann überaus nützlich und eröffnet neben der Möglichkeit, z.B. Varianzen zu quantifizieren, auch Einsichten in prinzipielle Stärken bzw. Schwächen eines gegebenen Stichprobenplans (vgl. Abschnitte 4 und 5). Und dies geht offensichtlich weit über das hinaus, was bei einer Quotenauswahl - mangels angemessener mathematischer Modelle - über Genauigkeit und Leistungsgrenzen der Datenerhebungstechnik gesagt werden kann.

Ebenso wie eine Quotenstichprobe schloß sich die Erstellung einer Zufallsstichprobe über die Einwohnermeldedateien der Gemeinden von vornherein aus, und zwar allein schon aus Kosten- und Kapazitätsgründen. Da nämlich diese Art der Stichprobenerstellung für Untersuchungen der Größenordnung des ALLBUS 1980 nicht zur Routine der kommerziellen Markt- und Meinungsforschungsinstitute in der Bundesrepublik gehört, wären von der Projektgruppe sowohl überaus umfangreiche theoretische Vorarbeiten zum Stichprobenplan als auch zur Beschaffung und Bearbeitung der notwendigen Dateien zu leisten gewesen. Von der personellen Ausstattung und von den zur Verfügung stehenden Mitteln her war dies nicht möglich.

Konsequenterweise wurde nach den entsprechenden Vorüberlegungen entschieden, den bei den kommerziellen Markt- und Meinungsforschungsinstituten weit verbreiteten dreistufigen Stichprobenplan auf der Grundlage des "ADM³-Mastersample" einzusetzen. Von Bedeutung war bei dieser Entscheidung auch, daß die Mehrzahl der großen repräsentativen Querschnittsuntersuchungen in den Sozialwissenschaften in den letzten Jahren mit Hilfe dieses Stichprobenplans durchgeführt wurden und somit die Vergleichbarkeit von Resultaten aus diesen Studien mit solchen aus dem ALLBUS 1980 - wenn überhaupt möglich - erleichtert würde.

Abschnitt 3.2 wird im folgenden dazu dienen, einen kurzen Abriß des verwendeten Stichprobenplans zu geben; die Abschnitte 3.2.1 bis 3.2.3 werden die einzelnen Stufen der Datenerhebung im Detail erläutern.

3.2 Der Stichprobenplan des ALLBUS 1980 im Überblick

Der ALLBUS 1980 gehört zur Kategorie derjenigen nationenweiten Querschnittsuntersuchungen, bei denen es aus praktischen Gründen nicht möglich ist, die einzelnen Interviews beliebig regional streuen zu lassen (vgl. S.131). Will man solche Befragungen auf kleine Flächeneinheiten konzentrieren, bieten sich dazu in der Bundesrepublik und West-Berlin Stimmbezirke an:

Dieser Grundidee folgend bestand die erste Stufe des Stichprobenplans in der Ziehung einer Zufallsstichprobe von 630 Stimmbezirken⁴, welche die Bundesrepublik und West-Berlin in der Art eines Netzes überdeckten.

Da selbstverständlich nicht alle Personen der Grundgesamtheit in den ausgewählten Stimmbezirken = Primäreinheiten befragt werden konnten - dies wären mehr als 250 000 Personen gewesen -, wurden in der zweiten Stufe des Stichprobenplans aus jeder der ausgewählten Primäreinheiten zufällig acht bzw. sieben (zur genauen Spezifikation vgl. Abschnitt 3.2.2) Privathaushalte entnommen. Aus diesen sollte dann in der dritten Stufe des Stichprobenplans "vor Ort", falls vorhanden, genau eine Person der Grundgesamtheit zufällig ausgewählt werden.

Die Festlegungen der Stichprobengrößen auf der ersten und zweiten Stufe hatten zur Folge, daß 4 620 Haushaltsadressen von den Interviewern anzulaufen waren. Es kamen schließlich 2 955 auswertbare Befragungen zustande. Die Sollvorgabe war 3 000 Interviews und diese Stichprobengröße ist als ein Kompromiß zu sehen, der die Genauigkeitsansprüche in Relation setzt zu den verfügbaren Mitteln.

Zeitlich hatten die Stichprobenziehungen auf der zweiten Stufe und der dritten Stufe einen Abstand von etwa einem halben Jahr. Dies bedeutet, daß lediglich solche Personen aus der Grundgesamtheit keine Chance hatten, in die Untersuchung zu gelangen, die nach Abschluß der Ziehungen auf der zweiten Stufe einen Haushalt bezogen, der zur Zeit der Stichprobenziehung in der betreffenden Primäreinheit noch nicht existierte. Mithin stimmen die Auswahlgrundlage und die in Abschnitt 2 definierte Grundgesamtheit weitgehend überein. Dies trägt mit dazu bei, daß der Stichprobenplan von seiner ganzen Anlage her in hohem Maße der Untersuchung den geforderten Charakter einer Momentaufnahme verleiht.

3.2.1 Die Zufallsauswahl der Primäreinheiten

Fehlt für eine Umfrage im nationalen Maßstab eine zentrale Liste für die Grundgesamtheit, so ist es - auch international - die Regel, zunächst die Durchführung der Studie auf "repräsentativ" ausgewählte regional begrenzte Flächen (Primäreinheiten oder Klumpen) zu beschränken. Beim ALLBUS 1980 stützte sich diese Auswahltechnik auf Primäreinheiten in der Größenordnung von Stimmbezirken.

Als Datenbasis stand die Stimmbezirkseinteilung des Bundesgebiets bzw. West-Berlins zur Bundestagswahl vom 3.10.1976 bzw. zur Wahl des Berliner Abgeordnetenhauses vom 2.3.1975 zur Verfügung. Um sehr kleine Primäreinheiten zu vermeiden, wurden vor der Ziehung alle Stimmbezirke mit weniger als 400 Wahlberechtigten mit benachbarten Stimmbezirken innerhalb der Gemeinde bzw. des Kreises zu synthetischen Stimmbezirken zusammengefaßt.

Die Grundgesamtheit der ersten Stufe des Stichprobenplans bestand letztlich aus 49 380 Primäreinheiten, von denen also jede mindestens 400 Wahlberechtigte (Stand: 1976 bzw. 1975) aufwies. Diese war als Datei sortiert nach Bundesländern, Regierungsbezirken, Kreisen und Gemeindegrößenklassen nach Boustedt (Boustedt 1975), was den Effekt einer Schichtung durch Anordnung hatte (es lag der Gebietsstand vom 1.1.1977 zugrunde, vgl. Schaefer, 1979: 22). Zusätzlich wurde für jede Primäreinheit ein Schätzwert für die Anzahl der in ihr vorhandenen Privathaushalte konstruiert, das sog. Bedeutungsgewicht: Vor der Zusammenfassung zu synthetischen Stimmbezirken wurde für jeden Stimmbezirk ermittelt, wie groß in der zugehörigen Gemeinde (Gebietsstand 1.1.1977 !) die Anzahl der Privathaushalte im Jahre 1970 war und welchen Umfang in diesem Jahre dort die Wohnbevölkerung⁵ hatte - beide Daten wurden ggf. durch Fortschreibung auf den Gebietsstand 1.1.1977 aus den Daten der Volkszählung von 1970 gewonnen (vgl. Schaefer 1979: 22f.). Die erste dieser Zahlen sei mit HH (Gem/70), die zweite mit WBev (Gem/70) bezeichnet. Des weiteren standen zur Verfügung der Umfang der Wohnbevölkerung am 1.1.1977 (= WBev (Gem/77)), die Anzahl der Wahlberechtigten 1976 bzw. 1975 (s.o.) (= WBer(Gem)) und die Anzahl der Wahlberechtigten im Stimmbezirk selber (= WBer(Sti)). Das Bedeutungsgewicht s wurde mit diesen Notationen definiert zu

$$(3.1) \quad s = HH(\text{Gem}/70) * \frac{W\text{Bev}(\text{Gem}/77)}{W\text{Bev}(\text{Gem}/70)} * \frac{W\text{Ber}(\text{Sti})}{W\text{Ber}(\text{Gem})}$$

und das Bedeutungsgewicht einer durch Synthetisieren entstandenen Primäreinheit war die Summe der Bedeutungsgewichte der diese Primäreinheit definierenden einzelnen Stimmbezirke.

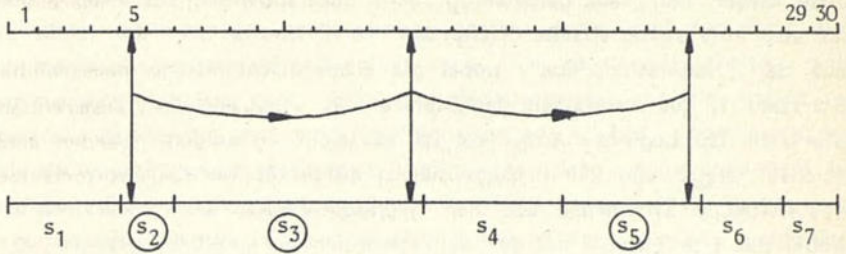
Eine solche Schätzung der Anzahl der Privathaushalte pro Primäreinheit ist selbstverständlich zu diskutieren. Da dazu jedoch der Zusammenhang mit der Art der Stichprobenziehung insgesamt hergestellt werden muß, wollen wir zunächst die Deskription des Stichprobenverfahrens fortsetzen und in Abschnitt 4.2 die Diskussion wiederaufnehmen.

Im Jahre 1978 wurde aus der so aufbereiteten Datei der Primäreinheiten eine Stichprobe der Größe 25 200 systematisch gezogen, und zwar in der Weise, daß die Wahrscheinlichkeit für eine bestimmte Einheit, in diese Stichprobe zu gelangen, proportional zu ihrem Bedeutungsgewicht s war.

Da es an dieser Stelle sicherlich unangemessen wäre, in voller Allgemeinheit diese Art der Ziehung abstrakt darzustellen, sei die Technik an einem einfachen Beispiel erläutert (der mathematisch interessierte Leser sei verwiesen an Des Raj 1968: 51f.):

Wir wollen annehmen, daß der gesamte Datensatz aus nur sieben Fällen besteht. Der i -te Fall möge das Bedeutungsgewicht s_i besitzen; im einzelnen: $s_1=4$, $s_2=2$, $s_3=9$, $s_4=5$, $s_5=5$, $s_6=3$, $s_7=2$. Die Aufgabe bestehe darin, eine systematische Stichprobe der Größe $n=3$ so zu ziehen, daß die Wahrscheinlichkeit für die i -te Einheit, in die Stichprobe zu gelangen, proportional zu s_i ist. Die Lösung ist aus Fig. 3.1 abzulesen. Jeder Einheit i ist dort eine ihrem Bedeutungsgewicht entsprechende Strecke s_i zugeordnet, so daß die Gesamtsumme der Strecken der Summe $S=30$ aller sieben Bedeutungsgewichte entspricht, und jede Einheit besetzt auf der Gesamtstrecke einen wohldefinierten Bereich. Man teilt nun diese Gesamtstrecke mit ihren 30 Einzelabschnitten in drei gleiche Teile zu je 10 Einzelabschnitten. Wählt man dann zufällig aus den ersten 10 Einzelabschnitten einen aus, etwa den 5ten, und fährt systematisch mit der Schrittweite 10 ($=S/3$!) fort, erhält also den 15ten und 25ten Einzelabschnitt, bedeutet dies die Ziehung einer Stichprobe der Größe Drei aus den sieben vorgegebenen Einheiten: zu den Einzelabschnitten 5, 15 und 25 gehören nämlich die Einheiten 2, 3 und 5.

Figur 3.1: Beispiel zur größenproportionalen Stichprobenziehung



Die Wahrscheinlichkeit, daß gerade die Stichprobe 2,3,5 ausgewählt wird, ist offensichtlich $4/10$ - diese Stichprobe erscheint nämlich genau für die Anfangsabschnitte 1 bis 4, die jeder für sich die Wahrscheinlichkeit $1/10$ tragen. Überhaupt möglich sind die fünf Stichproben 1,3,5 - Wahrscheinlichkeit $4/10$, 2,3,5 - Wahrscheinlichkeit $1/10$, 2,4,6 - Wahrscheinlichkeit $1/10$, 3,4,6 - Wahrscheinlichkeit $2/10$ und 3,4,7 - Wahrscheinlichkeit $2/10$. Man liest nun hieraus sofort ab, daß z.B. die Einheit 3 genau dann mit diesem Stichprobenplan gezogen wird, wenn eine von den vier Stichproben 1,3,5, 2,3,5, 3,4,6, 3,4,7 erscheint, also mit einer Wahrscheinlichkeit

$$4/10 + 1/10 + 2/10 + 2/10 = 9/10 = \frac{9 \cdot 3}{30} = \frac{s_3 \cdot n}{S}$$

Völlig analog stellt man fest, daß für die i -te Einheit allgemein die Ziehungswahrscheinlichkeit $s_i \cdot n / S$ beträgt, diese also proportional zum Bedeutungsgewicht s_i ist mit dem Proportionalitätsfaktor n/S .

Bis auf Modifikationen mehr technischer Natur⁶ wurde die Stichprobe von 25 200 Primäreinheiten wie im Beispiel dargestellt gezogen, so daß folglich jede der Einheiten mit der Wahrscheinlichkeit $s_i \cdot n / S$ behaftet war, in diese Stichprobe zu gelangen mit $n = 25\,200$ und $S = 22\,147\,308$. Im zweiten Arbeitsgang wurden danach 120 Unterstichproben der Größe 210 wie folgt definiert. Die erste Unterstichprobe bestand aus den Primäreinheiten mit den Nummern

1, 121, 241, ..., 24961, 25081, die zweite besaß die Einheiten
 2, 122, 242, ..., 24962, 25082, usw. und die 120te besaß die Einheiten
 120, 240, 360, ..., 25080, 25200.

Jede dieser 120, sich gegenseitig nicht überlappenden Unterstichproben ist also eine systematische Stichprobe - Schrittweite 120 - der Größe 210 aus der "Hauptstichprobe", wobei die erste dieser Stichproben mit der Startzahl 1, die zweite mit der Startzahl 2, usw. und die 120te mit der Startzahl 120 beginnt. Außerdem ist es leicht zu zeigen, daß ein jedes solches "Netz" von 210 Primäreinheiten aufgefaßt werden kann als eine systematische Stichprobe aus der Grundgesamtheit der Primäreinheiten, wobei das i -te Element mit der Wahrscheinlichkeit $s_i \cdot 210/S$, gezogen wird. Jedes Netz für sich genügt also bis auf einen anderen Proportionalitätsfaktor - $210/S$ anstelle von $25\ 200/S$ - der Ziehungsanforderung an die Hauptstichprobe, daß nämlich die Ziehungswahrscheinlichkeit für eine einzelne Einheit proportional zu ihrem Bedeutungsgewicht ist.

Die Hauptstichprobe zusammen mit den Kennungen für die einzelnen Netze wird auch "ADM-Mastersample" genannt.

Für den ALLBUS 1980 wurden drei dieser Netze herangezogen und die zugehörigen drei Startzahlen hatten den Abstand 21 bzw. 24 voneinander. Insgesamt kamen für die Befragung demnach 630 Primäreinheiten⁷ zum Einsatz und da die drei beteiligten Netze sich nicht überlappten, war die Wahrscheinlichkeit für irgendeine der Primäreinheiten mit dem Bedeutungsgewicht s_i in die Stichprobe der ersten Stufe zu gelangen, gleich $3 \cdot 210 \cdot s_i / S = 630 \cdot s_i / S$, so daß also nach wie vor die eingangs gestellte Proportionalitätsforderung in bezug auf das Bedeutungsgewicht erfüllt blieb.

Die dargestellte Stichprobentechnik geht zurück auf eine Idee von Hansen und Hurwitz (Hansen und Hurwitz 1943); sie wird insbesondere häufig dann angewendet, wenn eine Stichprobe aus einer Grundgesamtheit von Primäreinheiten, die jeweils unterschiedliche Größen aufweisen, zu ziehen ist und in nachfolgenden Arbeitsschritten aus den gezogenen Primäreinheiten nochmals Teilstichproben erhoben werden sollen (vgl. Kish 1965: 217ff.; Sudman 1976: 131ff.). Wesentlich ist dabei, daß man durch eine vorgegebene konstante Stichprobengröße innerhalb der Primäreinheiten bequem die Gesamtfallzahl im Griff behält und überdies - wie sich formal im Abschnitt 4 zeigen wird - mit dieser Technik Hochrechnungen vereinfachen kann.

Dem ersten dieser beiden Gesichtspunkte wurde beim ALLBUS 1980 in der Weise Rechnung getragen, daß für die 630 gezogenen Primäreinheiten jeweils 7 (in zwei Netzen) bzw. 8 (in einem Netz) Interviews als "Bruttovorgabe" festgesetzt wurden (vgl. GETAS 1980: 13).

Die Vereinfachung der Hochrechnung besteht darin, daß man die unterschiedlichen Auswahlwahrscheinlichkeiten für die Primäreinheiten nicht mehr berücksichtigen muß, wenn das Bedeutungsgewicht die Anzahl der Privathaushalte für jede der Primäreinheiten gut schätzt. Kommt noch eine konstante Stichprobengröße pro Primäreinheit dazu, kann zumindest bei lediglich zweistufigem Stichprobenplan gänzlich "ungewichtet" hochgerechnet werden. Die höhere Chance einer "großen" Primäreinheit, ausgewählt zu werden, wird nämlich gerade dadurch "kompensiert", daß die Anzahl der in ihr erhobenen Einheiten eben nicht ihrer Größe proportional ist, sondern jede Primäreinheit mit derselben Stichprobengröße bedacht wird. Einschränkend muß allerdings gesagt werden, daß dieses Argument nicht mehr als den intuitiven Hintergrund letztlich doch nur formal beschreibbarer Zusammenhänge beleuchten soll; es wird sich im Abschnitt 4 zeigen, daß nur unter ganz bestimmten Voraussetzungen der Vorteil vereinfachter Hochrechnung genutzt werden kann und daß konkret beim ALLBUS 1980 im Prinzip zwar diese Vereinfachung auch greift, die Konstruktion einer Gewichtung dadurch jedoch nicht entscheidend vereinfacht wird.

Selbstverständlich beeinflusst die Stichprobentechnik auf dieser ersten Stufe des Stichprobenplans die erreichbare Genauigkeit von Schätzungen durch Hochrechnung erheblich - eine detaillierte Diskussion findet sich in den Abschnitten 4 und 5. Typisch für systematisches, größenproportionales Ziehen ist dabei, daß Verteilungsparameter von Variablen, deren Verteilung zum Bedeutungsgewicht "paßt", mit großer Genauigkeit geschätzt werden können. Hat eine Variable jedoch nicht diese Eigenschaft, muß mit hoher und damit unerwünschter Variabilität der Schätzungen gerechnet werden.

Dies kann bedeuten, daß für bestimmte Variablen unter diesem Stichprobenplan Schätzungen mit einer höheren Varianz behaftet sind, als wenn sie aus einer durch uneingeschränkte Zufallsauswahl gewonnenen Stichprobe gleichen Umfangs hergeleitet worden wären. Zieht man zur Illustration das Beispiel dieses Abschnitts mit der Wertezuordnung $y_1=6$,

$y_2=7$, $y_3=3$, $y_4=2$, $y_5=6$, $y_6=1$, $y_7=1$ heran, zeigt sich, daß die Varianz des für das größenproportionale Ziehen unverzerrten Schätzers für den Mittelwert $\bar{Y} = 26/7$ etwa das Dreifache der Varianz des für uneingeschränkte Zufallsauswahl unverzerrten Schätzers (Stichprobengröße gleich drei!) $\hat{\bar{Y}}$ beträgt (vgl. auch Abschnitt 5, S.150).

Ein häufig benutztes Mittel, die Schätzerergebnisse aufgrund größenproportionalen, systematischen Ziehens zu verbessern, ist die zu Beginn des Abschnitts beschriebene Schichtung durch Anordnung (vgl. Sudman 1976: 134). Man möchte auf diese Weise erzwingen, daß, grob gesagt, die Stichprobengröße in jeder der Schichten proportional zur Schichtgröße in der Grundgesamtheit aller Primäreinheiten wird, daß also von den Stichprobengrößen in den einzelnen Schichten her die Situation des schichtweisen "proportionalen" Ziehens simuliert wird. Konkret ist es z.B. erwünscht, daß sich der Anteil der in Bayern gelegenen Primäreinheiten an der Gesamtheit aller Primäreinheiten in der gleichen Größenordnung in der Stichprobe wiederfindet.

Zwei Anmerkungen mögen dazu dienen, diese Technik der Schichtung durch Anordnung in den allgemeinen stichprobentheoretischen Rahmen zu stellen:

- Eine schichtweise gezogene Zufallsstichprobe - etwa von Primäreinheiten - unterscheidet sich wesentlich von einer Stichprobe, die systematisch nach schichtweiser Anordnung gezogen wurde. Typischerweise wird nämlich bei ersterem Vorgehen in den einzelnen Schichten unabhängig voneinander jeweils eine Zufallsstichprobe fest vorgegebener Größe erstellt, wohingegen bei systematischem Ziehen zum einen die Stichproben in den Schichten offensichtlich (stochastisch) voneinander abhängig und die Schichtstichprobengrößen zufällig sind. Konsequenterweise kann man bei Varianzberechnungen von Schätzungen nicht ohne weiteres auf Theoreme aus der Theorie des schichtweisen Ziehens zurückgreifen.

Wird im Zusammenhang mit dem ADM-Mastersample in der Literatur also von einer "mehrstufigen, geschichteten random-Stichprobe" gesprochen (vgl. Schaefer 1979: 93), so darf dies nicht zu dem Schluß verleiten, die Primäreinheiten seien im klassischen Sinne schichtweise gezogen worden.

- Will man über die Schichtung durch Anordnung Proportionalität erreichen, ist dies nur unter bestimmten Voraussetzungen möglich:

Wir wollen annehmen, daß die Summe aller Bedeutungsgewichte in der i-ten Schicht gleich S_i und, wie bereits zuvor benutzt, die Summe aller Bedeutungsgewichte gleich S ist. Es sei k_i die zufallsabhängige Anzahl derjenigen Primäreinheiten aus der Stichprobe der Größe k , die zur Schicht i gehören. Man überlegt sich nun leicht, daß der Erwartungswert von k_i gleich $k * (S_i/S)$ ist, also

$$Ek_i = k * (S_i/S).$$

Weist die Schicht i in der Grundgesamtheit aller Primäreinheiten N_i Elemente auf und ist die Größe dieser Grundgesamtheit gleich N , so bedeutet Proportionalität

$$Ek_i/k = N_i/N \text{ oder äquivalent } Ek_i = k * (N_i/N).$$

Letzteres ist jedoch gleichbedeutend damit, daß für jede der Schichten gilt $N_i/N = S_i/S$ oder äquivalent

$$S_i/N_i = S/N.$$

M.a.W. Proportionalität kann nur dann erwartet werden, wenn der Mittelwert der Bedeutungsgewichte in allen Schichten der Grundgesamtheit der Primäreinheiten derselbe ist.

Da der Autor bisher keinen Zugang zu dem entsprechenden Datenmaterial hatte, war es ihm nicht möglich zu prüfen, ob diese Bedingung für das ADM-Mastersample im wesentlichen erfüllt ist.

Selbstverständlich hat außer der Technik der Stichprobenziehung auch die Wahl gerade von Stimmbezirken als Primäreinheiten Einfluß auf die Präzision der Schätzungen von Parametern der Verteilung von Variablen. Genauer geht bei vorgegebener Variable unmittelbar in die Qualität der Schätzungen ein, ob ihre Verteilung innerhalb der Primäreinheiten im Schnitt von hinreichender Heterogenität ist ("Klumpeneffekt"; eine ausführlichere Betrachtung hierzu findet sich in Abschnitt 5). Bei der relativ geringen Größe von Stimmbezirken ist deshalb für viele Variablen eher mit einer gemessen an größeren Primäreinheiten geringeren Schätzerqualität zu rechnen. Dem ist jedoch entgegenzuhalten, daß für die Verwendung von Stimmbezirken zwei sehr gewichtige praxisbezogene Argumente sprechen:

- Zum einen wird durch eine einmalige zentrale Bearbeitung von Dateien wie der "Wahlbezirksdatei" (dies ist die Bezeichnung des Statistischen Bundesamtes für die Datei der Stimmbezirke) für mehrere Jahre Basismaterial für die Erhebung von Haushalts- bzw. Personenstichproben bereitgestellt. Es können so die Arbeitsschritte für die erste Stufe des Stichprobenplans in hohem Maße standardisiert werden, was wegen der folgenden Routinisierung Bearbeitungsfehler weitgehend eliminiert und nicht zuletzt kostengünstig ist.
- Zum anderen gestattet es die große Zahl der Primäreinheiten, vielen Instituten designgleiche, jedoch überschneidungsfreie Stichproben = Netze von 210 Primäreinheiten zur Verfügung zu stellen. Dies erleichtert dann vergleichende Untersuchungen über Studien verschiedener Institute. (Konkret wurden die 120 erstellten Netze auf 13 Institute verteilt; vgl. auch Schaefer 1979: 19ff.).

Zusammenfassend kann gesagt werden, daß der Stichprobenplan für die erste Erhebungsstufe des ALLBUS 1980 in theoretischer Hinsicht weitgehend erprobten Wegen folgt unter Einschluß bekannter Prozeduren zur Qualitätsverbesserung von Schätzungen; in bezug auf die Einarbeitung administrativer Gegebenheiten in der Bundesrepublik und West-Berlin sowie typischer Anforderungen, die sich aus den Arbeitsweisen kommerzieller Markt- und Meinungsforschungsinstitute ergeben, muß er als praxisgerecht bezeichnet werden.

3.2.2 Die Zufallsauswahl von Haushalten in den Primäreinheiten

Bei der Zufallsauswahl von Privathaushalten (zur Def. vgl. Anmerkung 1) in den 630 Primäreinheiten aus der ersten Stufe war zu bedenken, daß offenbar nur solche Haushalte von Interesse waren, in denen wenigstens eine mindestens 18jährige Person mit deutscher Staatsangehörigkeit lebte. Diese Art von Privathaushalten stellte also pro Primäreinheit die Grundgesamtheit auf der zweiten Stufe des Stichprobenplans dar, und daraus sollte in systematischer Zufallsauswahl eine Stichprobe gezogen werden. Es war dabei sicherzustellen, daß unter dem Gesichtspunkt "Momentaufnahme" das Adressenmaterial bis zur eigentlichen Personenbefragung nicht veraltet war.

Das für die Feldarbeit verantwortliche Institut GETAS (Gesellschaft für angewandte Sozialpsychologie, Bremen) ging diese in der kommerziellen Markt- und Meinungsforschung routinemäßig anfallende Problemstellung mit der folgenden Technik an.

In der zweiten Hälfte des Jahres 1979 wurden in den drei gezogenen Netzen von Primäreinheiten sog. Begehungsarbeiten durchgeführt, und zwar war in jedem der 630 Bezirke von Interviewern des Instituts nach festen Begehungsanweisungen eine "Wohnungstürklingel-Stichprobe" der Größe 200 in ein Adressenauflistungsblatt einzutragen.

Die weitgehend standardisierten Begehungsanweisungen regelten dabei im einzelnen, wie der Weg des Interviewers in der Primäreinheit verlaufen und auf welche Art die Adressenermittlung erfolgen sollte. Die Vielfalt der denkbaren örtlichen Gegebenheiten spiegelte sich selbstverständlich in dem erheblichen Umfang dieser Anleitungen und Vorschriften wider. Eine detaillierte Darstellung würde deshalb den hier gegebenen Rahmen sprengen; wir werden uns dementsprechend auf die Schilderung der zentralen Prozeduren beschränken:

Zur Festlegung des Weges, den der Interviewer bei der Auflistung gehen sollte, standen GETAS zum einen im Rahmen der Grundausstattung des ADM-Mastersample (vgl. Abschnitt 3.2.1, S.124) sog. Begehungsunterlagen, d.h. Beschreibungen und Abgrenzungen der betroffenen Stimmbezirke für solche Primäreinheiten zur Verfügung, die in Gemeinden mit mehr als 2000 Einwohnern⁸ lagen. Für jede dieser Primäreinheiten wurde im Institut eine feste Startadresse vorgegeben, von der ausgehend der Interviewer jeden Privathaushalt zu erfassen hatte, bis zur Erreichung des Soll von 200 Adressen. Der einzuschlagende Weg war von GETAS anhand der vorliegenden Straßenverzeichnisse im einzelnen vorgeschrieben. Verschiedene Varianten in den Vorgaben in bezug auf Kategorien wie "gerade Hausnummern", "ungerade Hausnummern", "nördliche Richtung", "südliche Richtung" wurden über die Primäreinheiten "gestreut". Bei der Wahl des Weges hatte der Interviewer also kaum Freiheiten.

Hatte GETAS für eine Primäreinheit keine Begehungsunterlagen zur Verfügung, war dem Interviewer vorgeschrieben, die Startadresse je nach den örtlichen Gegebenheiten festzulegen. Eine der Vorgaben bestand z.B. darin, in einer Straße mit einem bestimmten Anfangsbuchstaben des

Straßennamens mit der höchsten vorkommenden Hausnummer zu beginnen. Falls es keinen Straßennamen mit diesem Anfangsbuchstaben gab, mußte der im Alphabet folgende Buchstabe in gleicher Weise verwendet werden. Für den dann einzuschlagenden Weg wurden wiederum Anweisungen analog zu denjenigen ausgegeben, die auch beim Vorliegen von Begehungsunterlagen verwendet wurden.

Grundsätzlich wurde wiederum darauf geachtet, durch genaue Begehungsanweisungen dem Interviewer bei der Wahl des Weges möglichst wenig Freiheiten zu lassen. Die Anzahl der Einzelschriften ist daher beträchtlich, so daß sich allein schon aus Platzgründen deren weitere Darstellung, wie weiter oben bereits gesagt, an dieser Stelle verbietet. Der interessierte Leser sei in diesem Zusammenhang auf Schaefer (1979: 75ff.) verwiesen.

Vor Ort hatte der Interviewer die Erfassung der Privathaushalte zu operationalisieren über die Auflistung von Namen, die an den Wohnungstürklingeln standen. Bei einem Hochhaus war es ihm z.B. also nicht erlaubt, die Namen an der Haustürklingel zu verwenden. Des weiteren hatte er darauf zu achten, daß keine Adressen von Anstaltshaushalten, Gewerbebetrieben, Arztpraxen, Anwaltskanzleien usw. aufgenommen wurden, es sei denn, daß diese Adressen erkennbar mit einer Wohneinheit verbunden waren. In Zweifelsfällen sollte mit direkten Haushaltskontakten zur Nachfrage so sparsam wie möglich umgegangen werden, um u.a. zu verhindern, daß durch entstehendes Mißtrauen spätere Kontakte, die zu einem Interview führen sollten, zu Fehlversuchen wegen Verweigerung wurden.

Nach Abschluß der Begehungsarbeiten wurden von GETAS pro Primäreinheit aus den ersten hundert der 200 aufgelisteten Adressen in einem der drei Netze systematisch 8 Adressen, nämlich diejenigen mit den Nummern 1,13,25,...,85 entnommen; in den beiden anderen Netzen wurden auf die gleiche Weise im "14er Schritt" jeweils 7 Adressen gezogen, so daß insgesamt 4 620 Adressen als Brutto zur Ausgabe an die Interviewer bereitstanden (vgl. auch Abschnitt 3.2.1, S.125).

Im vorigen Abschnitt wurde die Stichprobengröße von 7 bzw. 8 Adressen pro Primäreinheit in den stichprobentheoretischen Rahmen der Ziehung auf der ersten Stufe des Stichprobenplans gestellt. Es sei hier darauf hingewiesen, daß diese Stichprobengrößen im Zusammenhang mit der Verwen-

dung von drei Netzen auch unter ökonomischen Gesichtspunkten überaus sinnvoll waren: Wären lediglich zwei Netze mit insgesamt 420 Primäreinheiten zum Einsatz gekommen, hätten bei dem zwingend notwendigen "Brutto" in der Größenordnung von 4 600 Adressen pro Primäreinheit mehr als 10 Adressen gezogen werden müssen. Dies hätte jedoch in der Regel den Einsatz von zwei Interviewern pro Primäreinheit gefordert, weil die Maximalzahl der Interviews pro Interviewer auf 7 festgelegt worden war. Nach Angaben von GETAS wäre dies teurer gewesen, als die tatsächlich gewählte Lösung mit drei Netzen, der Zuordnung eines Interviewers zu einer Primäreinheit und der Ausgabe von 7 bzw. 8 Adressen pro Primäreinheit. Dabei wurde bewußt in Kauf genommen, daß ein Interviewer 8 vollständige Interviews erreichen konnte. Die Erwartung, daß dies sehr selten geschehen würde, bestätigte sich deutlich, indem nur in 4 Primäreinheiten die Maximalzahl der Interviews überschritten wurde.

So wie bei der Festlegung der Stichprobengröße pro Primäreinheit Argumente ökonomischer Natur eine mitentscheidende Rolle spielten, wurde auch die soeben beschriebene Art der Stichprobenziehung durch Kostengründe stark beeinflußt: Es war ein Kompromiß zu finden zwischen der Anforderung, in der Primäreinheit systematisch eine Zufallsstichprobe von Privathaushalten zu ziehen, also eine starke räumliche Streuung der Adressen zuzulassen, und der Notwendigkeit, die Interviewerwege in Grenzen zu halten. Die von GETAS gezogenen Stichproben sind daher zwar technisch "systematisch" zustande gekommen und entsprechen allgemein üblichem Standard, weisen jedoch wegen der gewollten flächenmäßigen Beschränkung einen von der jeweiligen Primäreinheit abhängenden Klumpeneffekt auf. Dieser ist allerdings in seiner Wirkung auf die Genauigkeit von Schätzungen quantitativ kaum beschreibbar, dürfte aber, gemessen an den "nonsampling errors", eher gering sein für die meisten der interessierenden Variablen.

Die Operationalisierung der Stichprobe von Privathaushalten über eine Türklingel-Stichprobe hatte zur Konsequenz, daß Auswahlgrundlage und Grundgesamtheit pro Primäreinheit nicht deckungsgleich waren. So war wegen der Anweisung an die Interviewer, mit direkten Kontakten sparsam umzugehen, nicht auszuschließen, daß reine Ausländerhaushalte oder aber Untermieterhaushalte aufgelistet wurden, in denen keine Person der (Personen-)Grundgesamtheit lebte. Aus demselben Grunde mußte damit ge-

rechnet werden, daß sich hinter der nicht aufgelisteten Adresse etwa einer Anwaltskanzlei doch ein Privathaushalt der Grundgesamtheit verbarg. Des weiteren hatten Haushalte, die erst nach Adressenauflistung, jedoch vor der Personenbefragung, bezugsfertig wurden, keine Chance, in die Untersuchung zu gelangen.

Insgesamt haben solche Unterschiede zwischen Auswahlgrundlage und Grundgesamtheit stichprobentheoretisch gesehen stets einen negativen Effekt auf die Varianz von Schätzungen.

Seine Größenordnung ist aber auch hier nicht angebbbar; es ist zu vermuten, daß sie sehr gering ist. Hinzu kommt, daß sie in Relation gesetzt werden muß zu einem gewichtigen Vorteil, der mit dieser Art der Adressenermittlung, die gemeinhin Adress-Random genannt wird, verbunden ist: Bei der späteren Personenbefragung nimmt die Vorgabe einer festen Adresse dem Interviewer weitgehend die Freiheit zu entscheiden, ob er bei diesem Haushalt überhaupt einen Kontakt versuchen will. Ein solches unerwünschtes Verhalten ist zumindest denkbar, wenn die Nachteile des Adress-Random, nämlich Unschärfen bei der Erfassung der Grundgesamtheit und Aktualitätsverlust - hier etwa vier Monate -, umgangen werden, indem die Adresslistung und das eigentliche Interview gleichzeitig durchgeführt werden - "Standard-Random"-Technik.

Die Kontroverse Adress-Random versus Standard-Random, in die selbstverständlich in erheblichem Maße Aspekte der Interviewerkontrolle eingehen, ist gegenwärtig nicht entschieden. Im übrigen werden Überlegungen angestellt, durch völlig andere Techniken in den Primäreinheiten direkt Personenstichproben zu ziehen, um so die Nachteile beider Prozeduren zu vermeiden. Eine eingehendere Diskussion hierüber würde allerdings über den an dieser Stelle gegebenen Rahmen weit hinausgehen und muß daher unterbleiben.

3.2.3 Die Zufallsauswahl von Personen innerhalb der Haushalte

Zu Beginn des Befragungszeitraums wurde den einzelnen Interviewern ihr Adressenkontingent aus den 4620 Haushaltsadressen ("Klingeladressen") (vgl. voriger Abschnitt) zur Verfügung gestellt. Jeder dieser Adressen war zuvor ein Schema von Zufallszahlen angefügt worden, das es dem

Interviewer nach Auflistung aller Personen der Grundgesamtheit des angelaufenen Haushalts gestattete, aus diesem in uneingeschränkter Zufallsauswahl genau eine auszuwählen.

Dieses Verfahren sei anhand eines Beispiels erläutert.

Figur 3.2: Ausschnitt aus einer GETAS-Adressenliste

Ermitteln Sie die Befragungsperson, indem Sie alle Personen der Zielgruppe im Haushalt dem Alter nach geordnet – älteste Person zuerst – auflisten. Befragungsperson lt. Auswahlziffern bestimmen. (siehe Anleitung)															
ZIELPERSON (EN) IM HAUSHALT					ZAHL D. ZIEL-PERS.										
Alter		m/w		Alter	m/w		Auswahlziffern								
1.	75	W		6.			Haushaltsgröße (ZP)								
2.	42	W		7.			1	2	3	4	5	6	7	8	9
3.	39	W		8.			1	2	3	1	5	4	2	8	6
4.	19	W		9.			Auswahlziffer								
5.	18	W													

Die Figur 3.2 zeigt, daß der Interviewer im Zielhaushalt fünf Personen der Grundgesamtheit angetroffen hat und diese nach fallendem Alter unter Einschluß der Angabe "Geschlecht" von oben nach unten notierte. Unter die Zahlen von 1 bis 9 für die Anzahl der angetroffenen Zielpersonen (= Personen aus der Grundgesamtheit) sind jeweils (Zufalls-)Zahlen gedruckt, die stets kleiner oder gleich dem unmittelbar oberhalb stehenden Eintrag sind. Beispielhaft kann man sie sich entstanden denken aus acht unabhängig voneinander durchgeführten "Lottospielen" der Art "1 aus 2", "1 aus 3", ... und "1 aus 9" ("1 aus 1" ist trivial). Praktisch wurde die Prozedur maschinell ausgeführt. Wenn also im vorliegenden Beispiel der Interviewer diejenige Person befragt, deren Auflistungsnummer - als Zufallszahl - unter dem Eintrag für die Anzahl der Zielpersonen im Haushalt erscheint, hier demnach die fünfte, so hat er diese Person mit der Wahrscheinlichkeit 1/5 ausgewählt.

Allgemein hat bei diesem Verfahren eine Person in einem Haushalt mit i Personen der Grundgesamtheit die Auswahlchance $1/i$.

Es liegt hier natürlich die Frage nahe, warum eine solche relativ aufwendige Auswahlprozedur "vor Ort" nicht einfach dadurch ersetzt wird, daß man alle Personen der Grundgesamtheit aus dem Haushalt befragt. Insbesondere würde sich dann die Anzahl der auszugebenden Haushaltsadressen auf höchstens die Hälfte reduzieren.

Die Antwort ist sowohl inhaltlicher als auch statistischer Natur. Sie wurde 1949 von L. Kish formuliert für Umfragen, die 1946 bzw. 1947 in den U.S.A. unter der erwachsenen Bevölkerung durchgeführt wurden (Kish 1949):

"It was desired to take no more than one interview in any household, in order to obtain each interview before the respondent had a previous opportunity to discuss the questions. Furthermore, multiple interviews were thought to be statistically inefficient because of the expected correlation of attitudes within the household."

Die geschilderte Auswahlprozedur selber geht zurück auf eine Idee von L. Kish (Kish 1949) und heißt im deutschsprachigen Raum gemeinhin "Schwedenschlüssel". Sie wird in verschiedenen Spielarten, die sich in der jeweiligen Art der Zuordnung der Zufallszahlen niederschlagen, von der Mehrzahl der kommerziellen Markt- und Meinungsforschungsinstitute in der Bundesrepublik unter dem vorliegenden Stichprobenplan angewendet. Es sei jedoch betont, daß alle diese Verfahren letztlich nicht mehr als technische Varianten des Vorschlages von L. Kish sind.

4. Schätzaufbau im stichprobentheoretischen Modell

4.1. Zum Begriff der Repräsentativität eines Querschnitts

Der in den drei vorangegangenen Abschnitten vorgestellte Stichprobenplan wird bis auf geringe Modifikationen von den meisten der kommerziellen Markt- und Meinungsforschungsinstitute in der Bundesrepublik angewandt, wenn, wie beim ALLBUS 1980, eine Stichprobe bzw. ein "repräsentativer Querschnitt" aus der "erwachsenen Bevölkerung" zu erstellen ist.

Gegenstand dieses und der folgenden Abschnitte wird es sein, über die Deskription hinausgehend zu untersuchen, wie man sinnvollerweise aufgrund dieses Stichprobenplans Schätzungen von Verteilungsparametern berechnet und welche Qualität diese Schätzungen besitzen.

Es soll dabei scharf getrennt werden zwischen Überlegungen zur Leistungsfähigkeit des Stichprobenplans im Rahmen des stichprobentheoretischen Modells und Überlegungen zur Konstruktion von Schätzverfahren, wenn von einer durch sog. nicht-stichprobenneutrale Ausfälle betroffenen Stichprobe ausgegangen wird und deshalb die Diskussion sich nicht allein auf mathematische Argumente stützen kann. Dieser Abschnitt ist dementsprechend ausschließlich stichprobentheoretischer Natur und enthält - nach einer kurzen Diskussion des Begriffs der Repräsentativität - Aussagen zu prinzipiellen Stärken und Schwächen des vorliegenden Stichprobenplans.

Am Ende des Abschnitts über die Grundgesamtheit wurde bereits darauf hingewiesen, daß die Forderung nach einem "repräsentativen Querschnitt" von hoher Komplexität ist. Dies begründet sich darin, daß die Forderung nach "Repräsentativität" für sich alleine keinen Sinn macht, sondern nur im Zusammenhang mit den Zielen einer Untersuchung gesehen werden kann. Ist etwa eine Untersuchung orientiert an einer Gruppe von Variablen, die in einer Untergruppe der Zielpopulation kaum Variation aufweisen, so wird ein Stichprobenplan, der die Untergruppe gezielt mit einer geringen Fallzahl versieht, durchaus verlässliche Schätzergebnisse zur Folge haben, in diesem Sinne also repräsentativ sein. Derselbe Stichprobenplan kann jedoch zu völlig verzerrten Ergebnissen führen, wenn andere Variablen von Interesse sind, die bezüglich der Untergruppe gerade entgegengesetztes Verhalten haben. Von Repräsentativität kann dann also sicher nicht mehr gesprochen werden.

Sind, wie beim ALLBUS 1980, die Ziele der Untersuchung sehr weit gestreut, wäre es also eine schlechte Strategie, irgendwelche Gruppen der Grundgesamtheit in bezug auf ihre Abbildung in der Stichprobe von vornherein über den Stichprobenplan zu bevorzugen oder zu benachteiligen. Man wird vielmehr versuchen, der Stichprobe den Charakter einer "proportionalen" Abbildung der Population zu geben, oder doch zumindest die hinreichend gute Schätzbarkeit der Anteile von Subgruppen sicherzustellen.

Es wäre an dieser Stelle sicher unangemessen, eine mathematisch gehaltene Definition für einen repräsentativen Querschnitt zu geben (vgl. dazu Konijn 1973: 186f.). An welchem Begriff der Repräsentativität wir den Stichprobenplan zum ALLBUS 1980 messen wollen, sei daher wie folgt definiert:

Es sei eine Partition⁹ C der Grundgesamtheit gegeben, deren Elemente mit C bezeichnet seien; der Anteil von C an der Grundgesamtheit sei c - die Felder einer Kreuztabelle definieren z.B. stets in natürlicher Weise die Elemente einer Partition C . Wenn es dann für möglichst viele solcher, durch die Untersuchungsziele definierte, Partitionen C ein gemeinsames Schätzverfahren gibt, mit dem zumindest die größeren der Anteile c hinreichend genau geschätzt werden können, dann sei der aus dem Stichprobenplan resultierende Querschnitt repräsentativ genannt.

Diese Definition der Repräsentativität eines Querschnitts ist in ihren Anforderungen bewußt unscharf gehalten. Dies begründet sich zum einen darin, daß man stichprobentheoretisch zeigen kann, daß es keinen Stichprobenplan für Untersuchungen des Typs ALLBUS 1980 gibt, der garantieren würde, daß für alle, für die Untersuchungsziele theoretisch sinnvollen Partitionen C ein Schätzverfahren gefunden werden kann, das gleichmäßig über alle C 's mit vorgegebener relativer Genauigkeit Schätzungen der Anteile c liefert. Selbst mathematisch formulierte Abschwächungen einer oder aller dieser Forderungen führen nicht notwendig zu dem Schluß, daß dann ein repräsentativer Querschnitt herstellbar ist, ganz abgesehen von der großen Beliebigkeit, die solchen formalen Abschwächungen stets anhaftet.

Zum anderen würde nach allgemeinem Sprachgebrauch der Forschungspraxis in den Sozialwissenschaften einem Stichprobenplan, der die Anforderungen der gegebenen Definition erfüllt, kaum die Repräsentativität abgesprochen werden können.

Einzuwenden wäre lediglich, daß nicht unmittelbar einsichtig ist, wie mit Hilfe eines solchermaßen repräsentativen Querschnitts die Qualität von Schlußfolgerungen sichergestellt ist, die aus komplexeren statistischen Verfahren gewonnen werden, Verfahren also, die über Schätzungen von Mittelwerten weit hinausgehen. Ohne auf grundlegende Modelldiskussionen eingehen zu können (vgl. Saerndal 1978), sei hier folgendes angemerkt:

Zunächst bewirkt die Forderung, daß viele verschiedene Partitionen in ihren Einzelanteilen vermöge eines gemeinsamen Schätzverfahrens gut schätzbar sind, daß die Stichprobe in der Tat einem "proportionalen" Abbild der Grundgesamtheit nahekommte. Die Wahrscheinlichkeiten nämlich, mit denen die einzelnen Einheiten der Grundgesamtheit in die Stichprobe gelangen, können unter diesen Bedingungen nur wenig schwanken - was bekanntlich "Proportionalität" nach sich zieht. Die Begründung ist einfach. Würden die Auswahlwahrscheinlichkeiten zu einem Stichprobenplan einen größeren Wertebereich abdecken, wäre in aller Regel ein nicht vernachlässigbarer Teil der Grundgesamtheit mit besonders kleinen Auswahlwahrscheinlichkeiten behaftet. Selbst wenn es nun ein "gleichmäßig gutes" Schätzverfahren für viele Partitionen C gäbe, würde dieses ebenso in aller Regel den Anteil gerade der Elemente mit den besonders kleinen Auswahlwahrscheinlichkeiten schlecht schätzen, weil davon in der Stichprobe nur sehr wenige vertreten sein können. Dieser Querschnitt wäre also in bezug auf eine Partition C , die als Teil die Menge der Einheiten mit besonders kleinen Inklusionswahrscheinlichkeiten enthält, nicht repräsentativ.

Ein repräsentativer Querschnitt im Sinne obiger Definition ist also seiner "Proportionalität" wegen im Prinzip eine geeignete Datenbasis, um insbesondere statistische Analysen, die von mehrdimensionalen Kontingenztafeln ausgehen, durchzuführen. Zwei Einschränkungen sind jedoch zu beachten. Zum einen kann die Eigenschaft eines Querschnitts, "maßstabsgetreu" zu sein, erheblich beeinträchtigt werden durch das Auftreten sog. nicht-stichprobenneutraler Ausfälle. Im Abschnitt 9 über Gewichtungseffekte finden sich dazu grundsätzliche Bemerkungen. Zum anderen beeinflussen unter Umständen unterschiedliche Inklusionswahrscheinlichkeiten, selbst bei nur geringer Variation die Prüfverteilungen bestimmter Prüfgrößen vom χ^2 -Typ (vgl. Rao und Scott 1981), so daß also Inferenzschlüsse erschwert werden.

Für statistische Analyseverfahren, die mit den Werten von Variablen arbeiten, etwa solche im Rahmen linearer Modelle, ist zumindest, was den Fehler bei Mittelwertbildungen angeht, ein repräsentativer Querschnitt im Sinne der Definition ebenfalls eine geeignete Datenbasis. Es kann nämlich gezeigt werden, daß sich die geforderte gute Schätzbarkeit von Anteilen von Partitionen überträgt auf die Schätzung von Mittelwerten ganz allge-

mein. Wie jedoch unterschiedliche Auswahlwahrscheinlichkeiten und Einflüsse durch nicht-stichprobenneutrale Ausfälle bei Inferenzschlüssen mit Hilfe dieser Verfahren zu berücksichtigen sind, ist außerordentlich stark vom Einzelfall abhängig und kann an dieser Stelle nicht weiter diskutiert werden.

Insgesamt stellt also ein repräsentativer Querschnitt im Sinne der Definition Datenmaterial bereit, das auch den spezifischen Anforderungen, die sich aus sozialwissenschaftlichen Fragestellungen ergeben, weitgehend genügt. Zu untersuchen ist, ob der Stichprobenplan zum ALLBUS 1980 zu einem repräsentativen Querschnitt führte, zumindest im Rahmen des stichprobentheoretischen Ziehungsmodells.

Wir werden diese Fragestellung insofern positiv beantworten können, als ein Schätzverfahren angebbar ist, über dessen prinzipielle Stärken und Schwächen in bezug auf die mittleren quadratischen Abweichungen von den "wahren" Werten allgemeine Aussagen mathematischer Natur möglich sind.

4.2 Proportionalität, Auswahlwahrscheinlichkeiten und Schätzverfahren

Im vorigen Abschnitt zeigte sich, daß Repräsentativität eine Stichprobe in die Nähe eines proportionalen Abbildes der Grundgesamtheit bringen kann. Zentral für die Repräsentativität ist dabei der Begriff des gemeinsamen "guten" Schätzverfahrens. Bevor wir zeigen, daß ein solches in der Tat für den Stichprobenplan des ALLBUS 1980 mit gewissen Einschränkungen existiert, sei jedoch demonstriert, daß für eine wichtige Variable die Stichprobe nicht "proportional" ist.

Zu untersuchen ist dazu die Variabilität der Auswahlwahrscheinlichkeiten Π für die Einheiten (= Personen) U der Grundgesamtheit:

Es sei mit U eine beliebige Person der Grundgesamtheit bezeichnet. Diese möge in einer Primäreinheit PE wohnen und in einem Haushalt H leben. Die (unbekannte!) Anzahl der Privathaushalte zum Zeitpunkt der Erhebung in PE sei ξ (strenggenommen ist ξ die Anzahl derjenigen Privathaushalte in PE , die wenigstens eine Person der Grundgesamtheit enthalten) und in H mögen sich h Personen der Grundgesamtheit befinden. Das Be-

deutungsgewicht von PE sei s (vgl. Abschnitt 3.2.1), und die Stichprobengröße für die Haushaltsstichprobe in PE sei m . Legt man die Ziehungsprozeduren aus den Abschnitten 3.2.1 bis 3.2.3 zugrunde, sagt die Stichprobentheorie, daß die Wahrscheinlichkeit Π für die Einheit U , in die Stichprobe zu gelangen, genau

$$(4.1) \quad \Pi = \frac{630 * s}{S} * \frac{m}{\tilde{s}} * \frac{1}{h} \quad \text{ist.}$$

(Zur Bedeutung von S siehe Abschnitt 3.2.1).

Π variiert offenbar von Primäreinheit zu Primäreinheit in Abhängigkeit von den Größen s , \tilde{s} und m sowie von Person zu Person in der Größe $1/h$.

Da man in der Regel zumindest "Brutto" pro Primäreinheit im wesentlichen dieselbe Stichprobengröße m verwendet - wie auch beim ALLBUS 1980 - ist auf der Ebene der Primäreinheiten der Quotient s/\tilde{s} von besonderem Interesse. Dieser wird nichts zur Variation von Π beitragen, wenn das Bedeutungsgewicht s jeweils die Anzahl \tilde{s} der Privathaushalte pro PE gut schätzt (vgl. Abschnitt 3.2.1, S.125), wenn also $s/\tilde{s} \sim 1$. Es sind jedoch durchaus realistische Strukturmerkmale von Gemeinden angebar, unter denen für bestimmte Primäreinheiten der Quotient s/\tilde{s} deutlich von Eins abweicht: Angenommen, eine Gemeinde hatte 1970 10 000 und 1977 20 000 Einwohner, davon 13 000 Wahlberechtigte. 1970 möge die Anzahl der Privathaushalte 5 000 gewesen sein. In dieser Gemeinde liege - etwa in der "alten" Bausubstanz - eine Primäreinheit mit 1 000 Wahlberechtigten (1976) und $\tilde{s} = 500$ Privathaushalten, derselben Anzahl wie 1970. Das Bedeutungsgewicht beträgt unter diesen Gegebenheiten nach Formel (3.1), Abschnitt 3.2.1,

$$s = 5\,000 * 2 * (1\,000/13\,000) \cong 769, \text{ also } s/\tilde{s} \cong 769/500 = 1.54,$$

überschätzt demnach \tilde{s} erheblich.

Selbstverständlich sind auch andere Konstellationen innerhalb einer Gemeinde denkbar, die ebenfalls zu Abweichungen der Schätzung der Anzahl der Privathaushalte von dem tatsächlich gegebenen Wert \tilde{s} in bestimmten Primäreinheiten führen. Die Diskussion dazu soll hier nicht vertieft werden, und das soeben gegebene Beispiel möge als Problemhinweis insofern genügen, als praktisch die Abweichungen des Wertes s von \tilde{s} ignoriert werden müssen. Denn nur über eine Totalerhebung aller Privathaushalte in einer Primäreinheit ließe sich der aktuelle Wert des zugehörigen \tilde{s} an-

geben und so in (3.1) berücksichtigen. Dies ist in der Regel nicht durchführbar. Im übrigen sind dem Autor weder Untersuchungen über die Quantität der durch die Variation von s/\bar{s} hervorgerufenen Variation von Π bekannt, noch scheint es Untersuchungen zu geben, die ausschließen, daß die möglichen Fehlschätzungen von s u.U. die "Repräsentativität" der Netze von Primäreinheiten beeinträchtigen.

Im folgenden wird also stets davon ausgegangen, daß eine Person der Grundgesamtheit, die in der Primäreinheit PE in einem Haushalt H lebt, der h Personen der Grundgesamtheit enthält, mit der Wahrscheinlichkeit

$$(4.2) \quad \Pi = \frac{630}{S} * \frac{m}{h}$$

in die Stichprobe gelangt.

Folgt aus diesem Ansatz unmittelbar, daß jeder Privathaushalt dieselbe Chance hat, in die Stichprobe zu gelangen, so wird die Variation von Π auf Personenebene allein durch den Faktor $1/h$ bestimmt. Dies macht den Stichprobenplan in bezug auf Partitionen, die maßgeblich durch die Variable "Haushaltsgröße" bestimmt werden, "disproportional"!

Man sieht dies unmittelbar ein, wenn man die Erwartungswerte der zufälligen Größen n_i = Anzahl der Personen in der Stichprobe, die in einem Haushalt mit genau i Zielpersonen leben, berechnet.

Es ist nämlich

$$E n_i = \frac{630 * m}{S} * \frac{N_i}{i}, \text{ also } E [n_i/630 * m] = \frac{N}{S} * \frac{1}{i} * \frac{N_i}{N}, \text{ wobei } m$$

die zunächst als konstant angenommene Haushaltsstichprobengröße pro Primäreinheit ist (s.o.), N den Umfang der Grundgesamtheit bezeichnet und N_i das Analogon zu n_i in bezug auf die gesamte Zielpopulation darstellt.¹⁰

Offensichtlich ist $E[n_i/630 * m]$ nicht proportional zu N_i/N , dem entsprechenden Anteil in der Grundgesamtheit. M.a.W., im Schnitt ist z.B. der Anteil der Personen in der Stichprobe, die in einem Haushalt der reduzierten Haushaltsgröße¹¹ Eins leben, rund doppelt so groß wie N_i/N und der Anteil der Personen, die in einem Haushalt der reduzierten Haus-

haltsgröße Drei leben, beträgt im Schnitt rund $2/3$ der entsprechenden Größe aus der Grundgesamtheit. "Proportionalität", d.h. in diesem Falle erwartungstreue Schätzung von N_i/N , ist also nur durch "Gewichtung" von $n_i/630 * m$ mit $i * (S/N)$ zu erreichen.

Trotz des gewollt stichprobentheoretischen Charakters dieses Abschnitts sei kurz das soeben gezeigte - theoretische - Ergebnis an der Stichprobe geprüft. Man entnimmt einer Sonderauswertung des Mikrozensus 1980,¹² daß $N_1 = 6\,946\,600$ und $N_2 = 20\,005\,200$ mit $N = 44\,045\,500$.

Damit wird $(N_1/N) * 100 = 15.8\%$, $(N_2/N) * 100 = 45.4\%$. Die entsprechenden Prozentzahlen aus der Stichprobe sind 25.6% und 56.9% . Die erwarteten Prozentzahlen ergeben sich zu 31.4% und 45.2% . Die Abweichungen sind offenbar beträchtlich und nicht allein dadurch zu erklären, daß nach Abschluß der Feldarbeit die Anzahl der realisierten Interviews pro Primäreinheit schwankte, so daß die idealisierende Annahme eines konstanten m nicht erfüllt ist. Vielmehr muß vermutet werden, daß insbesondere für den zu hohen Anteil der Personen in Haushalten der reduzierten Haushaltsgröße zwei Ausfälle bei derjenigen Gruppe von Personen verantwortlich sind, die in den Haushalten der übrigen reduzierten Haushaltsgrößen leben.

Insgesamt ist also zu berücksichtigen, daß sich allein schon aus theoretischen Gründen in der Stichprobe kein "maßstabsgetreues" Bild der Variablen "reduzierte Haushaltsgröße" ergeben kann und daß zusätzlich dieser Effekt stark überlagert wird von systematischen Verzerrungen, die über das Feldgeschehen in die endgültige Stichprobe eingehen. Dieser Gedanke wird im Abschnitt 9 über Gewichtungseffekte wieder aufgenommen.

Bei den Überlegungen zum Erwartungswert des Anteils derjenigen Personen in der Stichprobe, die in einem Haushalt der reduzierten Haushaltsgröße i leben, zeigte sich, daß - theoretisch - der Ausdruck $\hat{N}_i = i * (S/N) * [n_i/630 * m]$ den "wahren" Anteil N_i/N erwartungstreu schätzt. Schreibt man \hat{N}_i in folgender Weise um,

$$\hat{N}_i = \frac{1}{N} * \sum_j^n \frac{S * i}{630 * m} * y_j ,$$

wobei n die Stichprobengröße ist, m die als konstant angenommene Anzahl der pro Primäreinheit erhobenen Haushalte bezeichnet und y_j genau dann den Wert Eins hat, wenn die j -te Person in der Stichprobe in einem Haushalt der reduzierten Haushaltsgröße i lebt - ansonsten ist $y_j = 0$ -, dann werden die "Ausprägungen" y_j offensichtlich gewichtet mit den Inversen der in Formel (4.2) angegebenen Auswahlwahrscheinlichkeiten. Im übrigen sieht man sofort, daß die Summe \hat{N}_i über die gewichteten Ausprägungen eine erwartungstreue Schätzung von N_i , also dem "Total" ist.

Die Ausdrücke für \hat{N}_i bzw. für \hat{N}_i sind Spezialfälle von sog. Horvitz-Thompson-Schätzern (Horvitz und Thompson 1952), die ganz allgemein den Mittelwert bzw. das "Total" einer Variablen - also etwa das Durchschnittseinkommen bzw. das Gesamteinkommen in einer Population - in intuitiv sehr naheliegender Weise erwartungstreu schätzen, indem die Ausprägungen der Variablen bei der Summation mit den Inversen der Auswahlwahrscheinlichkeiten multipliziert werden.

So ist z.B. auch das arithmetische Mittel $\sum_{i=1}^n x_i/n$ einer Variablen x gebildet

auf der Basis einer in uneingeschränkter Zufallsauswahl ohne Zurücklegen gezogenen Stichprobe der Größe n aus einer Population der Größe N , ein Horvitz-Thompson-Schätzer: Die Wahrscheinlichkeit für ein Element der Population, in die Stichprobe zu gelangen, ist n/N und offensichtlich gilt

$$\sum_{i=1}^n x_i/n = \left[\sum_{i=1}^n (n/N)^{-1} * x_i \right] / N .$$

Es soll nun zum Abschluß dieses Abschnitts gezeigt werden, daß, ausgehend von Horvitz-Thompson-Schätzern im Rahmen des vorgegebenen stichprobentheoretischen Modells, ein Schätzverfahren allgemein angegeben werden kann - also insbesondere auch für die Schätzung von Häufigkeiten -, das der Repräsentativitätsforderung insofern zumindest annähernd genügt, als prinzipiell Aussagen zur Schätzgenauigkeit möglich sind:

Der vorgegebene Stichprobenplan ist dadurch charakterisiert, daß mit einer Wahrscheinlichkeit proportional zum jeweiligen Bedeutungsgewicht s_i $l = 630$ Primäreinheiten gezogen wurden, daß aus der i -ten gezogenen Primäreinheit eine Zufallsstichprobe von Haushalten der Größe m_i ausgewählt wurde und daß im j -ten dieser Haushalte der reduzierten Größe h_{ij} genau eine Zielperson zu befragen war. Gibt man eine beliebige Variable Y vor, deren Ausprägung für den Befragten im j -ten Haushalt und der

i-ten Primäreinheit in der Stichprobe mit y_{ij} bezeichnet sei, so schreibt sich der Horvitz-Thompson-Schätzer für die Summe aller Ausprägungen von Y wie bereits erläutert in der allgemeinen Form

$$(4.3) \quad \hat{Y} = \sum_i \frac{S}{1 * m_i} * \sum_j^{m_i} h_{ij} * y_{ij} ,$$

vgl. auch Formeln (4.1); (4.2).

Die in der Forschungspraxis der Sozialwissenschaften überaus wichtige Mittelwertschätzung = Prozentschätzung, wenn Y eine dichotome "Zählvariable" ist, ergibt sich sofort zu $\hat{\bar{Y}} = \hat{Y}/N$, wenn N wieder den Umfang der Grundgesamtheit bezeichnet.

Genauigkeitsbetrachtungen zu den Schätzverfahren \hat{Y} bzw. $\hat{\bar{Y}}$ anzustellen wäre jedoch dem tatsächlichen Gebrauch von "Gewichten" - hier $(S/1 * m_i) * h_{ij}$ - in den großen Programmpaketen wie SPSS oder OSIRIS nicht angemessen. Dort wird gewichtet stets so prozentuiert, daß nicht durch N, sondern durch die Summe der Ausprägungen der Gewichtsvariablen dividiert wird. Dies geschieht bekannterweise vor allem, weil bei Prozentuierungen innerhalb von Subgruppen der Stichprobe der Umfang der Subgruppen in der Grundgesamtheit in der Regel nicht bekannt ist und so als Nenner der Prozentuierung nicht zur Verfügung steht.

Dementsprechend werden im folgenden die Eigenschaften des Schätzverfahrens

$$(4.4) \quad \hat{\bar{Y}} = \hat{Y} / \sum_i \frac{S}{1 * m_i} * M_i \quad \text{untersucht,}$$

wobei $M_i = \sum_j^{m_i} h_{ij}$ = Anzahl der in den zu kontaktierenden Haushalten in

der i-ten Primäreinheit der Stichprobe vorhandenen Zielpersonen. Man beachte, daß der Nenner von $\hat{\bar{Y}}$ zufallsabhängig ist.

Der Nenner von $\hat{\bar{Y}}$ ist im übrigen eine sehr plausible Schätzung für die Größe der Grundgesamtheit N . Faßt man nämlich M_i/m_i als Schätzung für die durchschnittliche Haushaltsgröße in der i -ten Primäreinheit auf und unterstellt, daß M_i/m_i nur unwesentlich von der durchschnittlichen reduzierten Haushaltsgröße N/S in der Grundgesamtheit abweicht, dann wird der Nenner von $\hat{\bar{Y}}$ zu

$$\sum_i (S/i) * N/S = N !$$

$\hat{\bar{Y}}$ ist kein erwartungstreus Schätzverfahren, d.h. es gilt nicht für alle Variablen Y , daß $E \hat{\bar{Y}} = \bar{Y}$ = Populationsmittelwert; deshalb interessiert bei Genauigkeitsuntersuchungen in erster Linie die mittlere quadratische Abweichung $E (\hat{\bar{Y}} - \bar{Y})^2$, die bekanntlich gleich der Varianz von $\hat{\bar{Y}}$ ist für $E \hat{\bar{Y}} = \bar{Y}$. Aus tieferliegenden stichprobentheoretischen Gründen kann $E (\hat{\bar{Y}} - \bar{Y})^2$ jedoch nur unter vereinfachenden Annahmen über den Stichprobenplan und dann auch lediglich in erster Näherung berechnet werden.

Im einzelnen wollen wir im folgenden annehmen:

- Die Ziehung der 630 Primäreinheiten kann als eine Ziehung mit Zurücklegen aufgefaßt werden. Es wird also theoretisch zugelassen, daß Duplikate in der Stichprobe der Primäreinheiten auftreten.
- Die Zuordnung der Haushaltsstichprobengrößen m_i soll vor der Ziehung der Primäreinheiten für alle Primäreinheiten erfolgt sein. Praktisch wurde diese Prozedur von GETAS nach der Ziehung für die drei eingesetzten Netze von Primäreinheiten durchgeführt, so daß tatsächlich die Allokation der m_i zufallsabhängig war.
- Die Ziehung von Haushalten innerhalb der Primäreinheiten wird als uneingeschränkt zufällig ohne Zurücklegen aufgefaßt. Man vergleiche hierzu die Diskussion in Abschnitt 3.2.2.

Es ist in diesem Rahmen nicht möglich, diese Vereinfachungen stichprobentheoretisch zu werten. Gleiches gilt für die mathematischen Begründungen zur Approximation von $E (\hat{\bar{Y}} - \bar{Y})^2$.

5. Die mittlere quadratische Abweichung der Schätzfunktion $\hat{\bar{Y}}$

In diesem Abschnitt wird zunächst eine formale Darstellung einer ersten Näherung für $E(\hat{\bar{Y}} - \bar{Y})^2$ gegeben; es schließt sich eine Diskussion darüber an, unter welchen allgemeinen Bedingungen die mittlere quadratische Abweichung klein wird. In Art eines Beispiels wird danach die Größe von $E(\hat{\bar{Y}} - \bar{Y})^2$ nach oben hin abgeschätzt, wenn \bar{Y} der Anteil in der Grundgesamtheit der Kohorte der Männer im Alter von 30 bis 39 Jahren ist.

Es sei wieder Y eine beliebige Variable, deren Mittelwert \bar{Y} mit Hilfe von $\hat{\bar{Y}}$ geschätzt werden soll. $\hat{\bar{Y}}$ kann dabei aufgefaßt werden als eine Funktion der überhaupt möglichen endlich vielen Stichproben ξ , also $\hat{\bar{Y}} = \hat{\bar{Y}}(\xi)$; die Variation von $\hat{\bar{Y}}$ wird demnach ausschließlich durch die Variation der Stichproben ξ gesteuert. Deshalb kann man schreiben:

$$(5.1) \quad E(\hat{\bar{Y}} - \bar{Y})^2 = \sum_{\xi} (\hat{\bar{Y}}(\xi) - \bar{Y})^2 P(\xi),$$

wobei $P(\xi)$ die Wahrscheinlichkeit bezeichnet, mit der eine Stichprobe ξ erscheint und die Summation sich über alle möglichen (Personen-)Stichproben erstreckt. Die mittlere quadratische Abweichung ist also ein gewichtetes Mittel der einzelnen quadrierten Abweichungen

$$\hat{\bar{Y}}(\xi) - \bar{Y} \quad (\sum_{\xi} P(\xi) = 1) \quad ^{13}.$$

Von einer interpretierbaren Darstellung von $E(\hat{\bar{Y}} - \bar{Y})^2$ wird man demnach erwarten, daß sich in ihr die bzgl. ξ variationserzeugenden Teile des Stichprobenplans widerspiegeln. Dies sind aber gerade die Stichprobenziehungen auf den einzelnen drei Stufen des Designs.

Eine Möglichkeit $E(\hat{\bar{Y}} - \bar{Y})^2$ unter den in Abschnitt 4, S.144, beschriebenen vereinfachenden Annahmen zu berechnen, ergibt sich durch Anwendung der sog. "taylorization"-Technik (vgl. dazu Des Raj 1968: 127). Etwas längliche Umformungen führen zu dem Resultat, daß in erster Näherung für die Schätzung des Anteils $p = \bar{Y}$ einer Kohorte (Y dichotom!) gilt:

$$(5.2) \quad E(\bar{Y} - p)^2 \doteq V_1 + V_2 + V_3 \quad \text{mit}$$

$$V_1 = \frac{S}{N^2 * I} * \sum_i \frac{1}{s_i} * (Y_i - p * H_i)^2 ,$$

$$V_2 = \frac{S}{N^2 * I} * \sum_i s_i * \left(\frac{1}{m_i} - \frac{1}{s_i} \right) * \frac{1}{s_i - 1} * \\ \sum_j \{ (Y_{ij} - p * h_{ij}) - D_i * \left(\frac{Y_i}{H_i} - p \right) \}^2 ,$$

$$V_3 = \frac{S}{N^2 * I} * \sum_i \frac{1}{m_i} * \sum_{h_{ij} > 1} h_{ij}^2 * p_{ij} * (1 - p_{ij}) .$$

Die verwendeten Symbole bedeuten im einzelnen:

\sum_i	Summation über alle L Primäreinheiten (L = 49 380)
Y_i	Anzahl der zur Kohorte gehörigen Personen der Grundgesamtheit in der i-ten Primäreinheit
H_i	Anzahl der Personen der Grundgesamtheit in der i-ten Primäreinheit
m_i	Stichprobengröße für die Haushaltsstichprobe in der i-ten Primäreinheit
\sum_j	Summation über alle Privathaushalte in der i-ten Primäreinheit, die eine Person der Grundgesamtheit enthalten
Y_{ij}	Anzahl der zur Kohorte gehörenden Personen der Grundgesamtheit im j-ten Privathaushalt der i-ten Primäreinheit
$\sum_{h_{ij} > 1}$	Summation über alle Privathaushalte in der i-ten Primäreinheit, die mehr als eine Person der Grundgesamtheit enthalten
h_{ij}	reduzierte Haushaltsgröße des j-ten Privathaushalts der i-ten Primäreinheit ($H_i = \sum_j h_{ij} !$) .

$D_i = H_i/s_i$ kann als "mittlere reduzierte Haushaltsgröße" in der i-ten Primäreinheit aufgefaßt werden. $p_{ij} = Y_{ij}/h_{ij}$ ist der Anteil der Kohorte im j-ten Privathaushalt der i-ten Primäreinheit. Die Größen S, N, L, s_i wurden weiter oben bereits eingeführt.

Die Angabe der mittleren quadratischen Abweichung lediglich für die Schätzung eines Anteils p begründet sich einzig und allein in dem Anliegen, den terminologischen Aufwand an dieser Stelle in Grenzen zu halten.

Im übrigen würde eine allgemeine Darstellung der Schätzung von \bar{Y} für die nachfolgende Diskussion keinen Gewinn bedeuten.

Wie oben gefordert, ergibt sich die Approximation von $E(\hat{\bar{Y}} - \bar{Y})^2$ in der Tat als eine Summe von drei Summanden V_1 , V_2 und V_3 , die ganz offensichtlich auf die drei Stufen des Stichprobenplans beziehbar sind. V_1 beschreibt die Abweichungen des Kohortenanteils pro Primäreinheit vom globalen Anteil p . V_2 leistet das Analoge für die Haushalte. V_3 beleuchtet die über die Personenauswahl im Haushalt erzeugte Variation in den Stichproben ξ .

Von besonderem Interesse ist selbstverständlich die Frage, wann V_1 , V_2 und V_3 klein werden. Allein unter dem Aspekt sehr großer Stichproben, also ohne zunächst Eigenschaften der betrachteten Variablen Y zu verwenden, sieht man für V_2 unmittelbar, daß für eine Totalerhebung in den beteiligten Primäreinheit, d.h. für $s_i = m_i$, dieser Anteil der mittleren quadratischen Abweichung verschwindet. Dasselbe gilt für den über die Personenauswahl im Haushalt entstandenen Anteil, wenn jeweils eine Haushaltstotalerhebung durchgeführt wird. Dies ist an dem Ausdruck für V_3 nicht unmittelbar zu sehen, kann aber leicht nachgerechnet werden. V_1 ist nicht zuletzt eine monoton fallende Funktion für wachsenden Stichprobenumfang l auf der Ebene der Primäreinheiten - bei festgehaltener Variable Y ! V_1 verliert allerdings seine Aussagekraft für zu große Werte von l , da die Approximation für $E(\hat{\bar{Y}} - \bar{Y})^2$ unter der vereinfachenden Annahme des Ziehens mit Zurücklegen hergeleitet wurde, die sicher für sehr große l nicht mehr zulässig ist.

Insgesamt gilt für alle drei Summanden, daß bei vorgegebener Variablen und wachsender Stichprobengröße auf einer der Stufen des Stichprobenplans die mittlere quadratische Abweichung abnimmt. Diese Aussage ist allerdings kaum überraschend und im Grunde für praktische Zwecke weniger interessant, als eine Aussage über das Verhalten von V_1 , V_2 und V_3 bei fest vorgegebenen Stichprobengrößen jedoch in Abhängigkeit von der Variablen Y .

Bei fester Stichprobengröße l hängt V_1 stark davon ab, ob $p \cdot H_i$ die Größe Y_i gut "schätzt", und zwar gleichmäßig über alle L Primäreinheiten. Schreibt man mit $p_i = Y_i/H_i$ die Größe V_1 um zu der Form

$$V_1 = \frac{S}{N^2 * I} * \sum_i H_i * D_i * (p_i - p)^2 ,$$

ist unmittelbar einsichtig, daß - äquivalent zur ersten Feststellung - V_1 wesentlich davon bestimmt wird, ob die Kohortenanteile p_i in den einzelnen Primäreinheiten mit dem Globalanteil übereinstimmen oder nicht. V_1 wird also sicher groß werden, wenn etwa für viele der Primäreinheiten mit großem H_i sich die p_i von p stark unterscheiden.

Die Variation der H_i ist im übrigen gemessen an derjenigen von D_i beträchtlich größer. Man kann dies mit Hilfe bekannter Daten über die 630 gezogenen Primäreinheiten leicht nachprüfen. Nimmt man nämlich an, daß die (bekannte) Anzahl der Wahlberechtigten in der i -ten Primäreinheit eine gute Näherung für H_i ist und nutzt aus, daß ohnehin hier $\hat{s}_i = s_i$ unterstellt wird - vgl. Abschnitt 4 - , so findet man für die vorliegenden 630 Fälle durch Nachrechnen: Mittelwert der $H_i = 972.35$, Standardabweichung der $H_i = 326.67$; Mittelwert der $D_i = 1.95$, Standardabweichung der $D_i = 0.33$; die vergleichsweise geringe Variation der D_i um den Wert 2 äußert sich auch in einer Regression in bezug auf die s_i (abhängig) und die H_i mit dem Regressionskoeffizienten 0.51, dem Achsenabschnitt 15.30 und der hohen "erklärten Varianz" von 0.75 ($\hat{s}_i = 0.51 * H_i + 15.30$); eine Folge hiervon ist der sehr kleine (Pearsonsche) Korrelationskoeffizient von -0.05 zwischen den H_i und den D_i .

Etwas vergrößernd gesagt, wird also V_1 in der Regel groß werden, wenn die Heterogenität der Primäreinheiten in bezug auf das (dichotome) Merkmal Y nicht hinreicht, um die einzelnen Klumpen = Primäreinheiten als Abbild der Gesamtheit auffassen zu können. Da letzteres eher für große als für kleine Klumpen zu unterstellen ist, wird hier deutlich, daß bei vorgegebener Auswahlmethode für eine gute Genauigkeit der Schätzung \hat{Y} ganz allgemein die Klumpen nicht zu klein sein dürfen. Genauigkeitsverluste durch solche Art von Primäreinheiten lassen sich auch nicht ohne weiteres durch einen höheren Stichprobenansatz I ausgleichen (vgl. auch die Ausführungen in Abschnitt 3.2.1, S.127). Im übrigen zeigt sich hier, daß auch aus stichprobentheoretischen Gründen die Zusammenfassungen sehr kleiner Stimmbezirke zu größeren Einheiten geboten war.

Ein einfaches hinreichendes Kriterium dafür, daß V_1 klein wird, ist die Proportionalität der Y_i und der s_i . D.h., wenn es eine Konstante A gibt, so daß für alle Primäreinheiten die dortige Anzahl Y_i der zur Kohorte gehörigen Zielperson sich schreiben läßt als $Y_i = A * s_i$, dann wird V_1 sehr klein. Man zeigt nämlich durch leichte Umformungen, daß unter der Voraussetzung $Y_i = A * s_i$ gilt

$$V_1 = \frac{S}{N^2 * 1} * A^2 \sum_i s_i * (1 - \frac{S}{N} * D_i)^2 .$$

Beachtet man nun, daß $S/N \sim \frac{1}{2}$ und die D_i - wie soeben gezeigt - eine relativ geringe Variation um 2 herum aufweisen, wird die Behauptung unmittelbar einsichtig.

Das Kriterium der Proportionalität von "Klumpentotal" Y_i und Bedeutungsgewicht s_i für einen kleinen Wert von V_1 ist eine ganz wesentliche Rechtfertigung dafür, auf der ersten Stufe eines Stichprobenplans die Primäreinheiten überhaupt proportional zu einem vorgegebenen Bedeutungsgewicht zu ziehen. Man möchte so ausnutzen, daß unter Einsatz der "Zusatzinformation", die in den s_i enthalten ist, auch bei kleinen Stichprobenumfängen der Wert von V_1 nahezu Null werden kann, wenn nur die Variation Y_i/s_i sehr klein ist.

Schätzt man unter $Y_i = A * s_i$ den Wert von V_1 nach oben ab, indem man $(1 - (S/N) * D_i)^2$ gleich 0,5² setzt, was also $D_i \sim 1$ bzw. $D_i \sim 3$ bedeutet, so so wird $V_1 \leq (22\ 147\ 308^2 / 44\ 045\ 500^2 * 630) * A^2 * 0.25 = A^2 * 0.0001003$, folglich $V_1^{1/2} \leq A * 0.010$; nimmt man für A den realistischen Wert 0.25 an, folgt $V_1^{1/2} \leq 0.0025$. Die "Standardabweichung" liegt also im Bereich eines Viertelpromille. Am Ende dieses Abschnitts wird sich zeigen, daß man unter ganz anderen Annahmen beim Schätzen des Anteils der Gruppe der Männer im Alter von 30 bis 39 Jahren zu ähnlichen Größenordnungen für eine Obergrenze von V_1 kommt. Dies stützt im übrigen die plausible Vermutung, daß für diese Kohorte die Variation von Y_i/s_i nicht sehr groß ist. Eine empirische Überprüfung hierzu aufgrund der (Personen)Stichprobe ist jedoch wegen der kleinen Fallzahl pro Primäreinheit nicht möglich.

Der Einsatz dieses sog. "pps-Ziehens" (pps = probability proportionate to size) garantiert also zumindest bei "geeignet" verteilten Variablen nahezu beliebig kleine Beiträge V_1 zur gesamten quadratischen Abweichung beim Schätzen eines Anteils p mit Hilfe von \hat{Y} . Es sei angemerkt, daß man eine solche Eigenschaft von "Varianzanteilen", die aus uneingeschränkter Zufallsauswahl (auf der ersten Stufe) stammen, nicht herleiten kann: Diese Art der Stichprobenziehung verzichtet bekanntlich auf den Einsatz von Zusatzinformation. Man wird demzufolge auch nicht erwarten können, daß

für Variablen Y , die mit ihrer Verteilung auf ein pps-Ziehen "passen", d.h. deren Verteilung sich bereits weitgehend über die (bekannte!) Zusatzinformation beschreiben läßt, dieselbe Schätzerqualität - wenn auch für die Schätzprozeduren selber keine Zusatzinformation verwendet wird - erreichbar ist, wie sie beim größtenproportionalen Ziehen und entsprechen der Schätzfunktion resultiert.

In Abschnitt 3.2.1 wurde ausgeführt, daß beim Schätzen eines Mittelwerts einer Variablen Y , die bzgl. der überhaupt möglichen Stichproben beim pps-Ziehen "unglücklich" verteilt ist, mit schlechten Schätzergebnissen gerechnet werden muß. Dies läßt sich jetzt wie folgt präzisieren: Zeichnet sich die dichotome Zählvariable Y durch eine hohe Variation der Werte Y_i/s_i aus,¹⁴ und schätzt man den Anteil p mit Hilfe von \bar{Y} , wird V_1 in der Regel sehr groß werden. Man sieht dies unmittelbar wegen

$$(5.3) \quad V_1 = \frac{S}{N^2 * I} * \sum_i s_i \left(\frac{Y_i}{s_i} - p * D_i \right)^2 ,$$

wenn man beachtet, daß die Variation der D_i gering ist. Im übrigen wurde das in Abschnitt 3.2.1 gegebene Beispiel zum größtenproportionalen Ziehen gerade so konstruiert, daß Y_i/s_i eine große Variation aufweist (vgl. S.122, 125f.), was in der Tat dazu führt, daß die Schätzung des Mittelwerts \bar{Y} aufgrund uneingeschränkter Zufallsauswahl derjenigen aus der pps-Auswahl überlegen war.

Die Anteile V_2 und V_3 an der mittleren quadratischen Abweichung $E(\hat{Y} - \bar{Y})^2$ sind weit weniger als V_1 über die Verteilung von Y zu beschreiben. Für V_2 kann allgemein lediglich gesagt werden, daß der Wert stark vom Verhalten der Varianz V_{2i} der Werte $Y_{ij} - p * h_{ij}$ in den Primäreinheiten abhängt. Es gilt nämlich

$$V_2 = (S/N^2 * I) \sum_i s_i (1/m_i - 1/s_i) * V_{2i}.$$

Also wird V_2 klein werden, wenn in vielen Primäreinheiten gilt, daß die Summe

$$\sum_j \{ h_{ij} \left(\frac{Y_{ij}}{h_{ij}} - p \right) - H_i \left(\frac{Y_i}{H_i} - p \right) \}^2 \quad \text{klein ist. Dies bedeutet}$$

wiederum, daß für fast alle Haushalte j in der i -ten Primäreinheit die mit der Haushaltsgröße h_{ij} multiplizierte Differenz des Anteils Y_{ij}/h_{ij} der Kohorte im Haushalt mit dem Gesamtanteil p im Schnitt nur geringfügig von der auf der Ebene der i -ten Primäreinheit analog gebildeten Größe abweicht. Vorhersagbar wird dies z.B. offenbar zutreffen, wenn Y die Zählvariable für die Kohorte der Männer in der Grundgesamtheit ist. Weiter verdeutlicht werden soll das Verhalten von V_2 in dem Beispiel am Ende dieses Abschnitts.

Beim Anteil V_3 sind nur die Produkte $p_{ij}(1-p_{ij})$ von der Variablen Y abhängig. Diese sind bekanntlich sehr klein für Werte p_{ij} in der Nähe von Null bzw. Eins und maximal - nämlich 0.25 - für $p_{ij} = \frac{1}{2}$. V_3 ist z.B. gleich Null, wenn die Kohorte gerade die Menge derjenigen Zielpersonen ist, die in Privathaushalten der reduzierten Haushaltsgröße zwei, also $h_{ij} = 2$, leben. Dies ist selbstverständlich ein extremes Beispiel, und es lassen sich Konstellationen solcher Art in fast beliebiger Anzahl angeben. Sehr viel informativer ist dagegen die Angabe der Größenordnung von V_3 , wenn mit \hat{Y} der Anteil p einer Kohorte geschätzt wird, die sich nicht auf Haushalte einer festen (reduzierten) Größe konzentriert. Das nachfolgende Beispiel wird deshalb außer für V_3 auch für V_1 und V_2 Abschätzungen nach oben angeben, wenn Y die Zählvariable für die Kohorte der männlichen Zielpersonen im Alter von 30 bis 39 Jahren ist.

Die fragliche Kohorte hat nach einer Sonderauszählung aus dem Mikrozensus 1980 durch das Statistische Bundesamt¹⁵ einen Anteil $p = 0.08$ an der Grundgesamtheit, und es ist plausibel anzunehmen, daß in jeder der L Primäreinheiten Personen aus der Kohorte vertreten sind, daß also alle Y_i einen von Null verschiedenen Wert besitzen, der sich mit unbekannten α_i schreiben läßt als $Y_i = \alpha_i * H_i$, wobei dann die Variation der α_i sich in abschätzbaren Grenzen hält:

Annahme: Es gelte $0.04 \leq \alpha_i \leq 0.12$; dies bedeutet, daß für den Anteil $p_i = Y_i/H_i$ der Kohorte in der i -ten Primäreinheit unterstellt wird, daß p_i den Globalanteil p relativ (!) um 50% überschreiten oder unterschreiten kann. Eine solche Annahme erscheint plausibel, ist aber empirisch bisher nicht nachgeprüft.

Annahme: $1.7 \leq D_i \leq 2.3$; es wird also unterstellt, daß die D_i im Bereich der aus der Stichprobe geschätzten Standardabweichung (s.o.) variieren.

Aus beiden Annahmen folgt, daß wegen

$$\underline{Y_i} = \alpha_i * (H_i/s_i) * s_i = \alpha_i * D_i * s_i = \underline{\lambda_i * s_i}, \text{ mit } \lambda_i = \alpha_i * D_i, \text{ gilt}$$

$$\underline{0.068 \leq \lambda_i \leq 0.276},$$

so daß $\underline{Y_i}$ nicht proportional zu s_i ist.

Formel (5.3), S.150, ergibt nun unmittelbar, daß

$$V_1 = \frac{S}{N^2 * 1} * \sum_i s_i * (\lambda_i - 0.08 * D_i)^2 \leq \frac{S^2}{N^2 * 1} * (0.276 - 0.08 * 1.7)^2,$$

$$\text{also } V_1 \leq \left(\frac{22147308}{44045500} \right)^2 * \frac{1}{630} * 0.0196 = 0.00000787,$$

$$\text{mithin } V_1^{\frac{1}{2}} \leq 0.0028.$$

Genau wie unter der globalen Annahme der Proportionalität von $\underline{Y_i}$ und s_i sowie der großzügigen Unterstellung, daß die D_i zwischen Eins und Drei schwanken (s. S.149), ergibt sich hier unter Nicht-Proportionalität und einer restriktiveren, jedoch realistischen Annahme über die Schwankungsbreite der D_i eine Größenordnung desjenigen "Fehlers", der allein aus der Stichprobenziehung der ersten Stufe herrührt, von höchstens etwa einem Viertelprozent.¹⁶ Um eine obere Schranke für V_2 angeben zu können, sei die weitere Annahme gemacht, daß in einem Privathaushalt höchstens eine Person der Kohorte vorkommt, daß also $\underline{Y_{ij}}$ entweder gleich Null oder gleich Eins ist.

Diese vereinfachende Annahme wird gestützt durch den Befund, daß in der Stichprobe des ALLBUS 1980 in rund 99% aller Haushalte, in denen überhaupt ein Kohortenmitglied enthalten war, genau ein Kohortenmitglied existierte (in 515 Haushalten war die Kohorte präsent, in 508 Haushalten lebte genau ein Mitglied, in 6 Haushalten lebten zwei Mitglieder und in einem Haushalt wurden drei Mitglieder angetroffen).

Unter dieser Annahme und durch Ausnutzen der Beziehung

$$\sum_j [(Y_{ij} - p * h_{ij}) - D_i \left(\frac{Y_i}{H_i} - p \right)]^2 \leq \frac{s_i}{4} * [\text{Max} - \text{Min}]^2, \text{ wobei}$$

Max = maximaler Wert der $Y_{ij} - p * h_{ij}$, und

Min = minimaler Wert der $Y_{ij} - p * h_{ij}$ (i ist hier fest!), erhält man unter der weiteren Vereinfachung $m_i = m = 4.5$, daß

$$V_2 \leq 0.0000638, \text{ also schließlich } V_2^{\frac{1}{2}} \leq 0.008.$$

Der allein aus der zweiten Stufe des Stichprobenplans resultierende "Fehler" liegt also unter den genannten Umständen in der Größenordnung von höchstens einem Prozent.

Zur Bestimmung der Größenordnung von V_3 sind keine weiteren Annahmen erforderlich. Beachtet man die Setzung $m_i = m = 4.5$ und berücksichtigt, daß Y_{ij} entweder Null oder Eins sein soll, können die Produkte $p_{ij} * (1 - p_{ij})$ nur noch die folgenden Werte annehmen:¹⁷

für die Haushalte mit $h_{ij} = 2$ ist $p_{ij}(1 - p_{ij})$ entweder 0 oder 1/4,

für die Haushalte mit $h_{ij} = 3$ ist $p_{ij}(1 - p_{ij})$ entweder 0 oder 2/9,

für die Haushalte mit $h_{ij} = 4$ ist $p_{ij}(1 - p_{ij})$ entweder 0 oder 3/16,

usw. Man stellt so leicht fest, daß für alle Haushalte mit $h_{ij} \geq 3$ die Werte von $p_{ij}(1 - p_{ij})$ höchsten 2/9 sein können. Setzt man ein, daß es 2 315 500 Haushalte mit $h_{ij} = 2$ und 878 000 Haushalte mit $h_{ij} \geq 3$ gibt, folgt sofort die Abschätzung

$$V_3 \leq 0.0000164, \text{ also } V_3^{\frac{1}{2}} \leq 0.0040.$$

Der allein durch die Anwendung des Schwedenschlüssels - oder der "Kish-table" - auf der dritten Stufe des Stichprobenplans (vgl. Abschnitt 3.2.3, S.132f.) verursachte "Fehler" liegt also unter den beschriebenen Annahmen in der Größenordnung von höchstens einem halben Prozent.

Interpretiert man die Wurzel aus der mittleren quadratischen Abweichung von \bar{Y} , also $\{E(\bar{Y} - \bar{Y})^2\}^{\frac{1}{2}}$, als "Standardabweichung", zeigt das Beispiel insgesamt, daß dieser Wert höchstens in der Größenordnung von einem Prozent liegt ($\{V_1 + V_2 + V_3\}^{\frac{1}{2}} \leq 0.0094$). Die Allgemeinheit dieser Aussage

scheint berechtigt trotz der Tatsache, daß $E(\bar{Y}^{\hat{}} - \bar{Y})^2$ durch die Summe $V_1 + V_2 + V_3$ lediglich in erster Näherung dargestellt wird und darüber hinaus die einzelnen Summanden unter einschränkenden Bedingungen (s. S.144) betrachtet wurden. Insbesondere weil letzteres zusammen mit der verwendeten mathematischen Technik bei den einzelnen Abschätzungen der Unterstellung einer "ungünstigen Situation" sehr nahe kommt, kann man also schließen, daß bei der Schätzung $\bar{Y}^{\hat{}}$ des Anteils p einer Kohorte, die sich bzgl. der Größenordnung von p und ihrer Verteilung über die einzelnen Primäreinheiten und Haushalte ähnlich verhält wie die Kohorte der männlichen Personen in der Grundgesamtheit im Alter von 30 bis 39 Jahren, der Variationskoeffizient¹⁸ von $\bar{Y}^{\hat{}}$ nicht größer ist als 12.5%.

6. Die Ergebnisse der Abschnitte 4 und 5 im Überblick

Bevor in den folgenden Abschnitten diskutiert wird, wie durch Modifikation von $\bar{Y}^{\hat{}}$ der außerhalb des probabilistischen Modells angesiedelten Ausfallproblematik Rechnung getragen werden kann, seien die Ergebnisse der beiden vorangegangenen Abschnitte zusammenfassend dargestellt.

Ausgangspunkt der Überlegungen war, ob mit Hilfe des Stichprobenplans zum ALLBUS 1980 theoretisch ein repräsentativer Querschnitt aus der Grundgesamtheit gezogen werden konnte. Es zeigte sich, daß bzgl. eines relativ weit gefaßten Repräsentativitätsbegriffs insofern eine befriedigende Antwort auf diese Fragestellung gegeben werden kann, als es ein "globales" Schätzverfahren für Häufigkeiten gibt, von dem prinzipielle Stärken und Schwächen angebbar sind. Dem steht auch nicht entgegen, daß der Stichprobenplan Kohorten, die typischerweise in kleinen Haushalten leben, gegenüber solchen in großen Haushalten systematisch in der Weise bevorzugt, daß die ersteren bezogen auf die Grundgesamtheit überproportional in der Stichprobe - im Schnitt - vertreten sind.

Zur Bewertung des Schätzverfahrens wurde die mittlere quadratische Abweichung des Schätzwertes vom "wahren" Wert herangezogen. Deren Berechnung war aus (mathematisch) technischen Gründen nur unter vereinfachenden Annahmen zur Ziehungstechnik (s. S.144) und im Resultat lediglich approximativ möglich. Auch die in Abschnitt 3.2.1, S.126, diskutierte "Schichtung durch Anordnung" wurde nicht berücksichtigt.

Dennoch kann - an dieser Stelle notwendig ohne weitere Diskussion - unterstellt werden, daß die hergeleiteten Ausdrücke für die mittlere quadratische Abweichung als Grundlage zur Beurteilung des Designs und des Schätzverfahrens gut geeignet sind. Im einzelnen ergab sich:

1. Alle drei Stufen des Stichprobenplans spiegeln sich jeweils mit einem Summanden in der mittleren quadratischen Abweichung des Schätzverfahrens (s. Formel (5.2)) wider. Der Anteil aus der ersten Stufe ist dabei weit besser in seinem numerischen Verhalten einzuordnen, als die beiden Anteile aus den nachfolgenden Stufen.

2. Entscheidend für geringe Genauigkeitsverluste auf der ersten Stufe des Designs ist eine möglichst kleine Variation über die Primäreinheiten des Produkts "Anteil der Kohorte in der Primäreinheit" * "durchschnittliche reduzierte Haushaltsgröße in der Primäreinheit".¹⁹

Anhand eines Beispiels läßt sich zeigen, daß auch bei nicht mehr vernachlässigbarer Variation dieses Produkts der Anteil der ersten Stufe nicht notwendig wesentlich größer ist als die Anteile an der mittleren quadratischen Abweichung der beiden folgenden Stufen.

Theoretisch kann die Ziehungsmethode auf der ersten Stufe in Verbindung mit dem vorgeschlagenen Schätzverfahren zu schlechteren Ergebnissen als eine uneingeschränkte Zufallsauswahl führen. Diese Gefahr besteht immer dann, wenn wegen der relativ geringen Größe der Primäreinheiten für bestimmte Variablen das obige Produkt große Variation aufweist.

3. Insbesondere für demographische Variablen ist es recht einfach - wie an einem Beispiel demonstriert wurde - im mathematischen Ziehungsmodell die Größenordnung der mittleren quadratischen Abweichung nach oben abzuschätzen.

7. Idealisierendes Stichprobenmodell und realisierte Stichprobe

In den beiden Abschnitten 4 und 5 wurde gezeigt, wie innerhalb eines idealisierenden mathematischen Ziehungsmodells ein auf den Stichprobenplan zugeschnittenes und in seiner Qualität beurteilbares Schätzverfahren für Anteilsschätzungen²⁰ begründet werden kann. Bewertet man dieses

Modell im Hinblick auf seine Realitätsnähe, ist insbesondere zu beachten, daß bei allen mathematischen Ableitungen die Einheiten der Untersuchung behandelt werden wie Kugeln in einer Urne, aus der durch Entnahme eine Stichprobe gezogen werden soll. Es wird also die uneingeschränkte Verfügbarkeit der Einheiten unterstellt, d.h. die Wahrscheinlichkeit für eine Einheit, in die Stichprobe zu gelangen, soll nur abhängen von der Technik der Stichprobenziehung und nicht von individuellen Eigenschaften, die z.B. Erreichbarkeit beschreiben.

Dieser Ansatz berücksichtigt damit offenbar nicht die praktische Erfahrung, daß ganz bestimmte Kohorten in vielen Querschnitten bei weitem nicht in der Anzahl vertreten sind, wie es die theoretischen Auswahlwahrscheinlichkeiten erwarten lassen. M.a.W., die Praxis zeigt, daß außer den Inklusionswahrscheinlichkeiten, die sich im klassischen mathematischen Ziehungsmodell ergeben, sehr wohl individuelle oder gruppenspezifische Eigenschaften die Chance einer Einheit, in der Stichprobe aufzutreten, beeinflussen. Ein stichprobentheoretisches Modell, in das solche empirischen Befunde eingearbeitet wären, sollte seiner Realitätsnähe wegen plausiblerweise zu verbesserten Schätzverfahren führen und ganz allgemein dazu beitragen, die Repräsentativität des jeweiligen Querschnitts mit größerer Sicherheit beurteilen zu können.

In der Tat gibt es in der stichprobentheoretischen Literatur einige Ansätze, um innerhalb mathematischer Modelle insbesondere die Tatsache der Antwortverweigerungen (nonresponse) zu beschreiben. Dabei wird jeder Einheit der Grundgesamtheit eine Antwortwahrscheinlichkeit zugeordnet, mit deren Inversen dann die erhobenen Ausprägungen gewichtet werden (vgl. z.B. Anderson 1979). Abgesehen davon, daß für komplexere Stichprobenpläne befriedigende Modelle dieser Art bisher nicht veröffentlicht wurden, birgt ganz allgemein die Angabe von Antwortwahrscheinlichkeiten erhebliche Probleme. Zunächst ist sicherlich eine starke Abhängigkeit dieser Größen vom Thema der Untersuchung anzunehmen. Des weiteren liegt eine hohe Beliebigkeit darin, zu entscheiden, von welchen individuellen Eigenschaften der Zielpersonen die Antwortwahrscheinlichkeiten abhängen sollen. Und nicht zuletzt wird deren empirische Überprüfung schwierig sein: der Kontextabhängigkeit wegen ist die Möglichkeit von Kontrolluntersuchungen beschränkt.

Leider steht dem berechtigten Wunsch nach einem der Realität besser angepaßten stichprobentheoretischen Modell für den ALLBUS 1980 entgegen, daß schon im Rahmen des klassischen Modells - und ein weitergehendes Modell müßte dies als Spezialfall enthalten! - vereinfachende Annahmen notwendig waren, um überhaupt tiefergehende Aussagen über die mittlere quadratische Abweichung des vorgeschlagenen Schätzverfahrens (4.4)

machen zu können. Die Verfeinerung des Modells würde also beim jetzigen Wissensstand lediglich dazu führen, daß theoretische Genauigkeitsaussagen entweder gar nicht möglich sind, oder aber in ihrem Wert stark relativiert würden durch sehr drastische simplifizierende Annahmen gerade bzgl. des Modells.

Dies bedeutet, daß bis auf weiteres zwischen den Problemkreisen des intuitiv begründeten und praktisch exekutierten Schätzeraufbaus per Gewichtung und des mehr deduktiven Schätzeraufbaus im Rahmen eines mathematisch noch beherrschbaren stichprobentheoretischen Ziehungsmodells nur so eine Verknüpfung hergestellt werden kann, indem untersucht wird, wie groß der "Abstand" der Schätzungen des ersteren Typs von denjenigen ist, die im vorigen Abschnitt dargestellt wurden. Daß also die "Modellseite" festgehalten und - wo möglich - an diese die "praktische Seite" herangeführt wird. Man definiert so letztlich die Qualität der Schätzung, die mit dem Ratio-Schätzer $\hat{\bar{Y}}$ in (4.4) erreichbar ist, als Bezugspunkt für das, was man über eine Gewichtung an Genauigkeit der Schätzung für Randverteilungen erreichen möchte.

Es sei angemerkt, daß der alleinige Bezug auf die "Basis" $\hat{\bar{Y}}$ zu sehen ist als Konsequenz aus dem im vorigen Abschnitt definierten Begriff der Repräsentativität (vgl. S.136). Von Interesse sind auch beim mehr intuitiv begründeten Aufbau einer Schätzung über "plausible" Gewichte lediglich solche Verfahren, die für möglichst viele theoretisch sinnvolle Partitionen brauchbar sind, also Gewichtungen, die einen universellen Charakter haben. Nicht untersucht werden deshalb z.B. spezielle Arten der Hochrechnung von Parteienpräferenzen, die mit Hilfe sog. Rückerinnerungsfragen versuchen, für die eine Variable, die aus der "Sonntagsfrage" resultiert, die Randverteilungen möglichst genau zu schätzen.

Nach Formel (4.4) hat der Schätzer $\hat{\bar{Y}}$ für den Anteil der Kohorte K die Gestalt

$$\hat{\bar{Y}} = \sum_i \frac{630}{m_i} \frac{1}{m_i} \sum_j h_{ij} * y_{ij} \bigg/ \sum_i \frac{630}{m_i} \frac{M_i}{m_i} ;$$

nimmt man jetzt die Stichprobengrößen m_i pro Primäreinheit als konstant an - die stichprobenneutralen Ausfälle verteilen sich gleichmäßig über alle Primäreinheiten ! - und erinnert sich, daß y_{ij} für den j-ten Befragten in der i-ten Primäreinheit genau dann gleich Eins ist, wenn dieser zur Kohorte gehört und Null sonst, so ergibt sich

$$(7.1) \quad \hat{\bar{Y}} = \sum_K h_j / \sum_{\xi} h_j ,$$

wobei $\sum_{\xi} h_j$ die Summe der zum j-ten Fall gehörigen reduzierten Haushaltsgrößen h_j über die gesamte Stichprobe ξ bedeutet; $\sum_K h_j$ ist die analoge Summation über alle zur Kohorte K gehörigen Fälle.

Die Schätzung $\hat{\bar{Y}}$ zeichnet sich offensichtlich dadurch aus, daß die verwendeten "Gewichte" h_j unabhängig sind von irgendwelchen Größen, die sich nur aus der - idealen - Gesamtstichprobe herleiten lassen und somit in aller Regel unbekannt sind. Diese Eigenschaft von $\hat{\bar{Y}}$ ist für den Stichprobenplan des ALLBUS 1980 typisch und sie erlaubt es - vgl. (9.4) -, in einfacher Weise Schätzungen, die aus der realisierten Stichprobe gewonnen wurden, zu vergleichen mit $\hat{\bar{Y}}$.

Selbstverständlich liegt es hier nahe zu fragen, warum man nicht direkt auch in bezug auf die realisierte Stichprobe Prozentanteile mit $\hat{\bar{Y}}$ aus (7.1) schätzt. Denn gänzlich unabhängig von stichprobentheoretischen Überlegungen ist es intuitiv einleuchtend, daß die Auswahl von genau einer Person pro Haushalt für die Stichprobe insofern disproportionierend wirkt, als Personen aus kleinen Haushalten relativ zu allen anderen Personen bei der Auswahl bevorzugt werden. Dieser Effekt bleibt natürlich bestehen, auch wenn die ideale Stichprobe durch Ausfälle "beschädigt" wurde. Es liegt also sehr nahe, der Disproportionalität vermöge einer Gewichtung mit der reduzierten Haushaltsgröße entgegenzuwirken. Zu beachten ist jedoch, daß so auch die Wirkung von Ausfällen mitgewichtet wird, daß also keineswegs die Schätzung notwendig verbessert wird. Zwei kleine Beispiele mögen dies erläutern:

Tabelle 7.1: Verteilung der Variablen "reduzierte Haushaltsgröße"

Kategorie	1	2	>2
Absolut	757	1680	518
Prozent (ungewichtet)	25.6	56.9	17.5
Prozent (gewichtet)	12.9	57.2	29.9
Mikrozensus 1980	15.8	45.4	38.8

Prozent (gewichtet) bedeutet Gewichtung²¹ mit der Variablen "reduzierte Haushaltsgröße", also Anwendung der Schätzung (7.1) auf die realisierte Stichprobe.

Offensichtlich verbessert diese Gewichtung die Schätzung für die beiden Randkohorten und ändert die Prozentzahl für die mittlere Kohorte, gemessen an der ungewichteten Berechnung nur geringfügig. Auffällig ist die Unterschätzung der Randkategorien bei gewichteter Berechnung, was sich in jeweils unterschiedlicher Überlagerung von Ausfall- und Gewichtungswirkung begründet:

Nach dem im Abschnitt 4.2, S.141, angegebenen Beispiel sind die erwarteten prozentualen Anteile für die Randkohorten 31.4% bzw. 23.4%. Dem stehen die zu niedrigen Anteile von 25.6% und 17.5% aus der realisierten Stichprobe gegenüber, in denen sich trotz des einzukalkulierenden Stichprobenfehlers klar ausdrückt, daß beide Kohorten von Ausfällen betroffen sind, und zwar weitaus stärker als die Mittelkategorie. Durch eine Gewichtung mit der reduzierten Haushaltsgröße wird nun der prozentuale Anteil derjenigen Personen in der Stichprobe, die in Haushalten der reduzierten Größe Eins leben, so vermindert, daß eine Unterschätzung des "wahren" Wertes resultiert. Umgekehrt reicht die durch eine solche Gewichtung bewirkte Erhöhung des prozentualen Anteils derjenigen Personen in der Stichprobe, die in Haushalten der reduzierten Haushaltsgröße von mindestens drei leben, nicht aus, um eine Unterschätzung zu vermeiden.

Für ein zweites Beispiel sei die Kohorte derjenigen Frauen in der Grundgesamtheit herangezogen, die wenigstens 70 Jahre alt sind. Ungewichtet beträgt deren Anteil in der Stichprobe 7.3%. Der Mikrozensus von 1980 weist einen Anteil von 8.9% aus, gewichtet mit der reduzierten Haushaltsgröße erhält man aus der Stichprobe 4.7%. Anders als im ersten Beispiel wird hier eine ungewichtet berechnet nahe am "wahren" Wert liegende Prozentzahl durch diese Art von Gewichtung deutlich verschlechtert.

Betrachtet man den erwarteten Anteil dieser Kohorte von rund 13% - in bezug auf die unbeschädigte Stichprobe -, so deuten die in der Stichprobe vorgefundenen 7.3% wiederum trotz einzukalkulierenden Stichprobenfehlers auf eine hohe Zahl von Ausfällen hin. Es kommt hinzu, daß die Kohorte sich - kontrolliert nach reduzierter Haushaltsgröße - wie folgt verteilt:

Tabelle 7.2: Weibliche Zielpersonen im Alter von wenigstens 70 Jahren nach reduzierter Haushaltsgröße (Mikrozensus 1980)

Kategorie	1	2	>2
Absolut (in 1000)	2026.4	1186.4	710.4
Prozent	51.6	30.2	18.1

Etwa die Hälfte der Kohorte lebt also in Haushalten der reduzierten Haushaltsgröße Eins, was natürlich bei der angewendeten Gewichtung zu einer Verminderung der ungewichtet berechneten Prozentzahl führen muß; dieser Effekt wird noch verstärkt durch den Umstand, daß in der Stichprobe 77.3% aller mindestens 70jährigen Frauen in Haushalten der Kategorie 1 leben.

Insgesamt wirken in diesem Beispiel die Tatsache hoher Ausfälle, die Gewichtung mit der reduzierten Haushaltsgröße und damit die Verteilung der Kohorte auf die einzelnen Haushaltsgrößenkategorien - wie bei der Schätzung der Randkategorie 1 des vorigen Beispiels - in ungünstiger Weise in derselben - falschen - Richtung, so daß ein sehr schlechtes Schätzergebnis die Folge ist.

Eine im Prinzip sinnvolle Gewichtung - wie hier diejenige mit der reduzierten Haushaltsgröße - kann also zu einer Verstärkung der disproportionierenden Wirkung von nicht-stichprobenneutralen Ausfällen führen. Um diesen Effekt abzuschwächen, wird häufig auf eine Technik zurückgegriffen, die eine Stichprobe an eine bekannte Verteilung "anpaßt"; es wird dabei durch eine entsprechende Gewichtung erzwungen, daß wenigstens für eine Partition die Anteile ihrer Elemente an der Grundgesamtheit "richtig" geschätzt werden. Im einzelnen wird dieses Gewichtungsverfahren im folgenden Abschnitt 8 dargestellt, der auch eine Beschreibung der im Datensatz zum ALLBUS 1980 enthaltenen beiden Gewichtungsvariablen enthält. Im abschließenden Abschnitt 9 werden die zugehörigen Gewichtungseffekte einer systematischen Betrachtung unterzogen.

8. Zwei Gewichtungsvorschläge für den ALLBUS 1980

Im Datensatz zum ALLBUS 1980, der beim Zentralarchiv für empirische Sozialforschung der Universität zu Köln für Interessenten bereitgehalten wird, sind zwei Gewichtsvariablen enthalten. In der Abfolge der Variablen-

nummern gelesen, wurde die erste bei GETAS, die zweite bei ZUMA konstruiert. Beide Variablen enthalten als Faktor die reduzierte Haushaltsgröße; damit wird - wie schon geschildert - der disproportionierenden Wirkung des Stichprobenplans Rechnung getragen. Die Tatsache, daß bei bestimmten Ausfallkonstellationen eine Gewichtung mit diesem Faktor allein zu verzerrten Schätzwerten führen kann, wird ebenfalls in beiden Variablen vermöge zusätzlicher sog. Redressementfaktoren berücksichtigt.

Bevor im Abschnitt 9 diskutiert wird, ob Gewichtungen dieses Typs zu einer Verbesserung der Schätzwerte führen können, sei in diesem Abschnitt zunächst die verwendete Redressement-Technik in allgemeiner Form dargestellt. Die nachfolgende Beschreibung des Aufbaus der zwei Gewichtsvariablen dient zur Erläuterung des allgemeinen Prinzips.

Liegt irgendeine Stichprobe ξ aus irgendeiner Grundgesamtheit G vor, und sind irgendwie "vernünftige" Gewichte w_i konstruiert worden, so daß der Anteil der Kohorte K and G geschätzt werden kann durch²²

$\sum_K w_i / \sum_{\xi} w_i$, ist eine Verbesserung dieser Schätzung im Prinzip möglich durch eine nachträgliche Schichtung der Stichprobe ξ und die Einführung von Schichtgewichten. Die Anteile der Schichten an der Grundgesamtheit sollen dabei möglichst gut bekannt sein. Werden die Schichten mit L_1, \dots, L_l und die zugehörigen Anteile an G mit $\text{Soll}(L_1), \dots, \text{Soll}(L_l)$ bezeichnet, so schreiben sich die mit dem Redressement bzgl. der Partition $C = \{L_1, \dots, L_l\}$ versehenen Gewichte w_i zu $\tilde{w}_i = \{[\text{Soll}(L_m) / \text{Ist}(L_m)] * w_i\}$, mit $\text{Ist}(L_m) = \sum_{L_m} w_j / \sum_{\xi} w_j$, wenn die i -te Einheit aus der Stichprobe ξ zur Schicht L_m gehört.

Da der Stichprobenplan zum ALLBUS 1980 dreistufig ist, also als eine Stichprobe aus Stimmbezirken, eine Stichprobe aus Haushalten bzw. eine Stichprobe aus Personen gelesen werden kann - natürlich bei "richtigem" Aggregationsniveau der betrachteten Variablen - ist auf allen diesen drei Stufen ein Redressement grundsätzlich möglich. Das GETAS-Redressement bezieht sich auf die Stufen zwei und drei:

Auf Haushaltsebene - also der Ebene der Gesamtheit derjenigen Privathaushalte in der Bundesrepublik und West-Berlin, in denen zum Zeitpunkt der Befragung wenigstens eine mindestens 18 Jahre alte Person mit deutscher Staatsangehörigkeit lebte, vgl. auch S.128 - wurde von GETAS eine

Partition = Kreuztabelle bzgl. der beiden Variablen Bundesländergruppen und politische Gemeindegrößenklasse zugrunde gelegt; diese ist in Tabelle 8.1 wiedergegeben.

Tabelle 8.1: Redressement - Tabelle (GETAS) auf Haushaltsebene

	Berl.	SH+ HH	NS+ HB	NRW	Hes	RhP+ Saar	BW	Bay	
- 5 000	--	0.90	1.74	0.26	1.86	2.54	1.41	8.80	Ist
	--	1.36	1.90	0.06	1.32	2.72	2.72	6.86	Soll
- 20 000	--	1.57	3.64	4.02	3.97	1.89	5.57	2.55	Ist
	--	1.08	3.56	3.87	3.75	1.97	4.97	4.54	Soll
-100 000	--	1.19	2.88	8.72	1.82	1.70	4.53	3.30	Ist
	--	0.96	3.55	9.92	1.76	1.77	4.38	2.10	Soll
-500 000	--	0.56	2.06	8.74	1.45	1.79	1.87	2.18	Ist
	--	0.80	2.08	8.15	1.14	1.26	1.81	1.93	Soll
500 000 u. mehr	2.67	2.44	1.96	5.48	0.61	--	0.78	2.54	Ist
	3.20	2.78	1.82	5.78	1.03	--	0.97	2.13	Soll

Kreuztabellierung von politischer Gemeindegrößenklasse und Bundesländergruppen

Auf Haushaltsebene wurden also 40 Schichten gebildet und pro Schicht wurde die Soll-Zahl (in Prozent!) der amtlichen Statistik entnommen. Die zugehörigen Ist-Zahlen wurden - ebenfalls in Prozent - ungewichtet aus der (Haushalts-) Stichprobe berechnet. Dies bedeutet, daß die unterschiedliche Anzahl der erfolgreichen Interviews in den einzelnen Primär-

Tabelle 8.2: Redressement - Tabelle (GETAS) auf Personenebene

		Berl.	SH+ HH	NS+ HB	NRW	Hes	RhPf+ Saar	BW	Bay	
♂	18 - 29	0.16	0.87	1.46	3.64	0.98	0.99	1.54	2.17	Ist
		0.33	0.69	1.28	2.93	0.98	0.81	1.65	1.89	Soll
	30 - 39	0.25	0.63	1.43	2.63	0.76	0.55	1.90	1.62	Ist
		0.34	0.73	1.25	2.77	0.90	0.72	1.51	1.73	Soll
	40 - 49	0.31	0.39	0.91	3.03	0.60	0.82	1.49	1.39	Ist
		0.24	0.56	1.06	2.67	0.75	0.67	1.23	1.45	Soll
	50 - 59	0.27	0.74	1.11	2.43	1.28	0.56	1.37	1.36	Ist
		0.26	0.56	1.00	2.20	0.71	0.62	1.09	1.38	Soll
	60 -	0.45	0.79	1.45	2.78	1.04	0.60	1.54	1.52	Ist
		0.88	1.31	2.19	4.43	1.49	1.26	2.24	2.86	Soll
♀	18 - 29	0.37	0.84	1.52	2.71	1.13	1.12	1.78	2.24	Ist
		0.36	1.11	1.40	3.00	1.03	0.86	1.71	1.94	Soll
	30 - 39	0.42	0.74	1.19	2.11	0.72	0.64	1.29	1.72	Ist
		0.39	0.78	1.35	3.01	1.01	0.76	1.68	1.89	Soll
	40 - 49	0.47	0.91	1.45	3.25	0.88	0.85	1.49	1.66	Ist
		0.22	0.53	0.83	2.52	0.78	0.67	1.28	1.44	Soll
	50 - 59	0.06	0.30	0.96	2.57	0.60	0.77	1.26	1.94	Ist
		0.15	0.36	0.70	1.61	0.52	0.44	0.77	0.96	Soll
	60 -	0.41	0.74	1.44	2.52	1.01	0.80	1.21	1.89	Ist
		0.42	0.81	1.31	2.76	0.97	0.81	1.42	1.81	Soll

einheiten vernachlässigt wird (vgl. auch S.157) und somit nicht in eine (Design-) Gewichtung auf Haushaltsebene eingeht.

Als "vernünftiges" Ausgangsgewicht, das die Rolle des w_i auf Personenebene übernimmt, ergab sich so das Produkt von Soll/Ist auf Haushaltsebene und der reduzierten Haushaltsgröße. Die Redressement-Partition bestand hier aus einer Kreuztabellierung von drei Variablen: Bundesländergruppen, Altersklassen und Geschlecht. Tabelle 8.2 zeigt die Kategorisierung dieser Variablen.

Auf Personenebene ergaben sich also 80 Schichten und pro Schicht wurde wiederum die Soll-Zahl (in Prozent) der amtlichen Statistik entnommen. Die Ist-Zahlen (in Prozent) errechneten sich, mit dem Ausgangsgewicht

$[Soll/Ist(Haushaltsebene)] * \text{red. Haushaltsgröße}$

gewichtet(!), unmittelbar aus der Stichprobe.

Insgesamt ist das GETAS-Gewicht also das Produkt der drei Faktoren Soll/Ist(Haushaltsebene), reduzierte Haushaltsgröße und Soll/Ist(Personenebene).

Das ZUMA-Gewicht wurde für Auswertungen solcher Art konstruiert, bei denen zwar gewichtet gerechnet werden soll, wo aber entweder eine Anpassung an Zensus-Zahlen unerwünscht ist, oder aber auf Haushalts- bzw. Personenebene andere Redressement-Partitionen als beim GETAS-Gewicht benutzt werden sollen. Im einzelnen besteht das ZUMA-Gewicht aus den folgenden drei Faktoren:

- einem Redressementfaktor auf der Ebene der Primäreinheiten,
- dem Inversen der Anzahl der erfolgreich Befragten in der Primäreinheit und
- der reduzierten Haushaltsgröße.

Dem Redressementfaktor auf der Ebene der Primäreinheiten liegt eine Kreuztabellierung der Variablen Stadtindex, Bedeutungsgewicht und Bundesländergruppen zugrunde; diese ist in Tabelle 8.3 wiedergegeben.

Tabelle 8.3: Realisierte und nicht realisierte Primäreinheiten nach Ländergruppen, kategorisiertem Bedeutungsgewicht und Stadt-Index

Stadt-index	Bedeutungsgewicht	SH,HH, HB,NS,B	NRW,He	BW,Rpf, Saar	Bay	Total
1	1 - 600	37 (2)	38 (3)	20 (1)	39 (1)	134 (7)
	600	4 (1)	2 (0)	7 (1)	6 (0)	19 (2)
2	1 - 600	11 (0)	24 (0)	29 (0)	16 (0)	80 (0)
	600	8 (1)	10 (2)	9 (0)	2 (0)	29 (3)
3	1 - 600	3 (2)	68 (0)	32 (0)	9 (0)	112 (2)
	600	2 (0)	23 (1)	7 (2)	1 (0)	33 (3)
4	1 - 600	12 (0)	15 (0)	10 (1)	13 (0)	50 (1)
	600	8 (0)	7 (0)	5 (0)	2 (0)	22 (0)
5	1 - 600	34 (4)	18 (2)	10 (0)	0 (0)	62 (6)
	600	24 (2)	16 (2)	3 (0)	18 (0)	61 (4)
Total		143 (12)	221 (10)	132 (5)	106 (1)	602 (28)

Der Eintrag in Klammern ist die Anzahl der nicht realisierten Primäreinheiten.

Zum Stadtindex vgl. Hoffmeyer-Zlotnik 1981; zum Bedeutungsgewicht sei auf S.121f. verwiesen.

Abweichend von den Tabellen 8.1 und 8.2 werden in Tabelle 8.3 absolute Zahlen für Soll und Ist ausgewiesen. Dies begründet sich darin, daß die Soll-Prozentzahlen unbekannt sind und somit Soll/Ist geschätzt werden muß, und zwar durch den Quotienten A/B, wobei

A = Anzahl der ursprünglich gegebenen Primäreinheiten in der Redressement-Zelle und

B = Anzahl derjenigen Primäreinheiten in der Redressementzelle, in denen mindestens ein Interview erfolgreich abgeschlossen wurde.

Im übrigen ist das "vernünftige" Gewicht, das dem Redressement auf der Ebene der Primäreinheiten sinnvollerweise zugrunde gelegt wird, $w_i = S/603 * s_i$, wenn s_i das Bedeutungsgewicht der i-ten Primäreinheit ist (vgl. S.124).

Insgesamt ist das ZUMA-Gewicht das Produkt der drei Faktoren Soll/Ist (Ebene der Primäreinheiten), 1/Anzahl der Interviews in der Primäreinheit und reduzierte Haushaltsgröße.

Beide Gewichte sind durch einen zusätzlichen Normierungsfaktor auf die Fallzahlen 3 014 (GETAS) bzw. 2 955 (ZUMA) standardisiert, d.h. die Summe der 2 955 Ausprägungen ist 3 014 bzw. 2 955.

9. Gewichtungseffekte

Zu Beginn des Abschnitts 7 wurde gezeigt, daß sich im stichprobentheoretischen Modell aus Abschnitt 4 die Schätzung \hat{Y}^a (vgl. (7.1)) für den Anteil einer Kohorte K an der Grundgesamtheit darstellen läßt über eine Gewichtung durch die reduzierten Haushaltsgrößen allein. In den nachfolgenden Beispielen wurde demonstriert, daß diese Gewichtung, angewandt auf die realisierte Stichprobe, sowohl zu Verbesserungen als auch zu erheblichen Verschlechterungen von Schätzwerten führen kann.

Gegenstand dieses Abschnittes ist es, die Systematik, aus der solche Gewichtungseffekte resultieren, darzustellen. Insbesondere wird der Einfluß von Redressementfaktoren untersucht werden. Zur Beurteilung der Qualität einer gewichtet berechneten Anteilsschätzung wird der angekündigte "Abstand" zur "Modellseite" - nämlich zu \hat{Y}^a - angegeben.

Vervollständigt man die Tabelle 7.1 durch diejenigen Prozentzahlen, die nach Gewichtung mit dem ZUMA-Gewicht bzw. mit dem GETAS-Gewicht berechnet werden, ergibt sich:

Tabelle 9.1: Verteilung der Variablen "reduzierte Haushaltsgröße"

Kategorie	1	2	>2
Absolut	757	1680	518
Prozent (ungewichtet)	25.6	56.9	17.5
Prozent (gewichtet mit red. Haushaltsgröße)	12.9	57.2	29.9
Prozent (GETAS-Gewicht)	16.2	58.0	25.8
Prozent (ZUMA-Gewicht)	12.9	57.6	29.6
Mikrozensus 1980	15.8	45.4	38.8

Vervollständigt man in gleicher Weise die Prozentzahlen für die Kohorte der mindestens 70jährigen Frauen und stellt diese in den Gesamtzusammenhang einer Tabelle für Alters- und Geschlechtskohorten, erhält man:

Tabelle 9.2: Altersklassen gegen Geschlecht

	18 - 29	30 - 39	40 - 49	50 - 59	60 - 69	≥ 70	
♂	9.3	8.9	9.6	7.1	6.6	4.5	u
	11.0	8.5	10.6	8.6	6.7	4.3	h
	10.7	10.4	8.7	5.8	6.6	4.3	hβ (G)
	11.6	8.8	10.9	8.6	6.7	4.0	hβ (Z)
	11.2	8.0	9.3	7.4	5.0	5.1	M (80)
♀	10.9	10.0	8.4	8.9	8.6	7.3	u
	11.1	9.3	9.3	9.3	6.5	4.6	h
	9.8	9.4	9.0	8.1	9.9	7.5	hβ (G)
	10.8	9.3	9.2	9.3	6.2	4.6	hβ (Z)
	10.8	8.1	9.3	9.2	7.7	8.9	M (80)

u: ungewichtet; h: gewichtet mit der reduzierten Haushaltsgröße; hβ (G): GETAS - Gewicht; hβ (Z): ZUMA - Gewicht; M (80): Mikrozensus 1980.

Sowohl in Tabelle 9.1 als auch in Tabelle 9.2 sind deutliche Gewichtungseffekte festzustellen. Zur Vorbereitung der Diskussion über die Systematik hinter diesen Effekten diene ein Beispiel, in dem die Gewichtungen die Prozentzahlen kaum verändern. In Frage 17 des Fragebogens geht es um Aussagen über das Verhalten von Beamten bei Behörden. Der Befragte sollte mit Hilfe einer 7-Punkte-Skala (1 = stimme überhaupt nicht zu, 7 = stimme voll und ganz zu) seine Meinung zu ganz bestimmten Aussagen ausdrücken. Eine dieser Aussagen war: "In Behörden wird man wie eine Nummer behandelt":

Tabelle 9.3: Zustimmung zur Aussage "In Behörden wird man wie eine Nummer behandelt"

Kategorie	1	2	3	4	5	6	7
Absolut	368	306	405	466	452	353	592
Prozent (ungewichtet)	12.5	10.4	13.8	15.8	15.4	12.0	20.1
Prozent rH	12.6	10.6	13.6	15.6	15.5	12.1	19.9
Prozent GG	12.9	10.7	13.8	15.3	16.1	11.6	19.7
Prozent ZG	12.4	10.5	13.9	15.2	16.4	11.5	20.1
Mittlere red. Haushaltsgröße	2.00	2.03	1.97	1.96	2.01	2.01	1.97

rH = reduzierte Haushaltsgröße; GG = GETAS-Gewicht; ZG = ZUMA-Gewicht

Beide in Abschnitt 8 vorgestellten Gewichte sind offenbar vom Typ $w = h * \beta$, wobei h die reduzierte Haushaltsgröße beschreibt und der Faktor β in sich die Soll/Ist-Anpassungen sowie die Standardisierungen auf die Zahlen 3 014 bzw. 2 955 vereinigt; β hängt somit beim GETAS-Gewicht auch von h ab. Um die Schreibweisen im folgenden zu vereinfachen, sei definiert:

Ist eine Kohorte K gegeben, deren Anteil an der Grundgesamtheit geschätzt werden soll, sei der bisher verwendete Ratioschätzer geschrieben als

$$(9.1) \quad \hat{p}(K|h\beta) = \frac{\sum_K h_i * \beta_i}{\sum_{\xi} h_i * \beta_i} \quad .^{23}$$

Um terminologisch in der Notation der einzelnen Schätzer die realisierte Stichprobe bzw. die "ideale" = unbeschädigte Stichprobe bequem kennzeichnen zu können, seien alle Größen, die sich über die letztere definieren, mit einem hochgestellten Kreis (°) versehen. So wird z.B. $\hat{\bar{Y}}^{\circ}$ (vgl. (7.1)) zu $\hat{P}(K^{\circ}|h)$, da ja hier die β_i konstant gleich Eins sind und nicht mehr explizit notiert werden müssen.

Weiter sei

$$\bar{h}(\xi) = \sum_{\xi} h_i/n, \quad \bar{h}(K) = \sum_K h_i/k, \quad \bar{\beta}(\xi) = \sum_{\xi} \beta_i/n, \quad \bar{\beta}(K) = \sum_K \beta_i/k,$$

$$\bar{h}\bar{\beta}(\xi) = \sum_{\xi} h_i * \beta_i/n \text{ und } \bar{h}\bar{\beta}(K) = \sum_K h_i * \beta_i/k,$$

wobei n = Stichprobengröße der realisierten Stichprobe ξ ,

und k = Anzahl der Personen aus der Kohorte K in der realisierten Stichprobe ξ .

Mit Hilfe dieser Schreibweisen wollen wir zunächst darstellen, wie Unterschiede im Resultat zwischen dem ohne Gewicht berechneten Schätzer $\hat{P}(K|\cdot) = k/n$ und $\hat{P}(K|h\beta)$ erklärbar sind.

Es gilt in erster Näherung

$$(9.2) \quad \hat{P}(K|\cdot) \doteq \frac{\bar{h}(\xi)}{\bar{h}(K)} * \frac{\bar{\beta}(\xi)}{\bar{\beta}(K)} * \exp \left\{ \frac{\bar{h}\bar{\beta}(\xi)}{\bar{h}(\xi)*\bar{\beta}(\xi)} - \frac{\bar{h}\bar{\beta}(K)}{\bar{h}(K)*\bar{\beta}(K)} \right\} * \hat{P}(K|h\beta);$$

der Faktor $\exp\{\dots\}$ ist lediglich ein Korrekturterm, der auch geschrieben werden kann als

$$\exp \left\{ [r(\xi, h, \beta) * \frac{\text{Std } h(\xi)}{\bar{h}(\xi)} * \frac{\text{Std } \beta(\xi)}{\bar{\beta}(\xi)}] - [r(K, h, \beta) * \frac{\text{Std } h(K)}{\bar{h}(K)} * \frac{\text{Std } \beta(K)}{\bar{\beta}(K)}] \right\};$$

dabei sei $r(\xi, h, \beta)$ der Pearsonsche Korrelationskoeffizient in bezug auf die beiden Variablen (!) h und β für die Fälle aus der Stichprobe ξ . $r(K, h, \beta)$ wird auf der Basis von K berechnet. $\text{Std } h(\xi)$ sei die Standardabweichung der Variablen h , wiederum bezogen auf die Fälle der Stichprobe ξ . Die Bedeutungen von $\text{Std } \beta(\xi)$, $\text{Std } h(K)$ und $\text{Std } \beta(K)$ sind entsprechend zu verstehen.

Man kann zeigen, daß die Darstellung in (9.2) in weiten Bereichen der Gleichheit beider Seiten sehr nahe kommt und der Korrekturterm $\exp\{\dots\}$ stets nahe bei Eins liegt. Daher ist die folgende Interpretation von (9.2) erlaubt:

Bezogen auf den ungewichtet berechneten Anteil $\hat{P}(K|\cdot)$ läßt sich die Wirkung der Gewichtung des Typs $h \cdot \beta$ allein schon vermöge der Mittelwerte ihrer Komponenten h und β in der realisierten Stichprobe ξ bzw. in der Kohorte K in der Stichprobe beschreiben. Entscheidend ist dabei, wie die Mittelwerte von h und β bzgl. der Kohorte abweichen von denjenigen, die sich auf die Gesamtstichprobe ξ beziehen.

Diese Interpretation ist selbstverständlich noch wenig hilfreich, wenn in prognostischer Weise die Frage beantwortet werden soll, in welchen Situationen sich denn überhaupt gewichtet und ungewichtet berechnete Prozentzahlen unterscheiden - ohne zunächst darauf einzugehen, welche dieser Zahlen die "richtigen" sind. Eine einfache Antwort kann es jedoch nicht geben, da sich die Einflüsse der Faktoren $\bar{h}(\xi)/\bar{h}(K)$ und $\bar{\beta}(\xi)/\bar{\beta}(K)$ in unterschiedlicher Weise überlagern können. Im einzelnen ergibt sich nämlich:

- a) $\bar{h}(\xi) > \bar{h}(K)$ und $\bar{\beta}(\xi) > \bar{\beta}(K)$ bedeutet in aller Regel²⁴,
daß $\hat{P}(K|\cdot)$ größer ist als $\hat{P}(K|h\beta)$.
- b) $\bar{h}(\xi) \leq \bar{h}(K)$ und $\bar{\beta}(\xi) \leq \bar{\beta}(K)$ bedeutet in aller Regel, daß $\hat{P}(K|\cdot)$
kleiner oder gleich $\hat{P}(K|h\beta)$ ist.
- c) $[\bar{h}(\xi) > \bar{h}(K)$ und $\bar{\beta}(\xi) \leq \bar{\beta}(K)]$ oder
 $[\bar{h}(\xi) \leq \bar{h}(K)$ und $\bar{\beta}(\xi) > \bar{\beta}(K)]$
kann beide möglichen Resultate nach sich ziehen.

Für die Konstellation a) diene die Kohorte der 30-39jährigen Frauen als Beispiel. Für das GETAS-Gewicht gilt in diesem Fall, daß $\bar{\beta}(\xi)/\bar{\beta}(K) = 1.0385$. Für den Quotienten der mittleren reduzierten Haushaltsgrößen erhält man $\bar{h}(\xi)/\bar{h}(K) = 1.0699$. Der Korrekturterm hat den Wert 0.9576. Setzt man in die rechte Seite von (9.2) noch für $\hat{P}(K|h\beta)$ den Wert von 9.4 ein, ist das Produkt dieser vier Faktoren gleich 10.0, dem Wert der ungewichtet berechneten Prozentzahl, die also in der Tat größer ist als die gewichtet berechnete.

Der Tabelle 9.3 entnimmt man, daß für die Kategorie 5 die mit dem ZUMA-Gewicht berechnete Prozentzahl 16.4 beträgt; sie ist also größer als die ungewichtet mit 15.4 berechnete Zahl. Dies ist ein Beispiel für die Konstellation b); es sind nämlich $\bar{h}(\beta)/\bar{h}(K) = 0.9900$, $\bar{\beta}(\xi)/\bar{\beta}(K) = 0.9434$ und der Korrekturfaktor = 1.0006 (vgl. auch Tabelle 9.5).

In der Konstellation c) wird der Gewichtungseffekt dadurch bestimmt, ob einer der beiden Faktoren überwiegt, oder ob sich beide in ihrer Wirkung gegenseitig aufheben. Für letzteres ist die Schätzung für den Anteil der Kohorte der mindestens 70jährigen Frauen ein gutes Beispiel, wenn man das GETAS-Gewicht heranzieht: es sind $\bar{h}(\xi)/\bar{h}(K) = 1.5794$, $\bar{\beta}(\xi)/\bar{\beta}(K) = 0.6429$ und der Korrekturfaktor $\exp\{\dots\} = 0.9671$. Die beiden fast identischen Schätzerergebnisse von 7.5 (gewichtet) und 7.3 (ungewichtet) finden sich in Tabelle 9.2.

Für den Fall, daß ein Faktor stark überwiegt, diene die Kohorte der 18-29jährigen Männer als Demonstrationsobjekt: für das ZUMA-Gewicht hat man $\bar{h}(\xi)/\bar{h}(K) = 0.8468$, $\bar{\beta}(\xi)/\bar{\beta}(K) = 0.9434$ und der Korrekturfaktor ist gleich 1.0014. In Tabelle 9.2 sieht man, daß der gewichtet berechnete Wert tatsächlich beträchtlich größer ist als der ungewichtet berechnete, daß sich also der Quotient der mittleren reduzierten Haushaltsgrößen als dominant erweist.

Besonders einfach sind im übrigen jetzt die numerischen Effekte der Gewichtung durch die reduzierte Haushaltsgröße allein erklärbar - wiederum gemessen an der ungewichteten Berechnung. Setzt man nämlich β_i konstant gleich Eins, verzichtet also auf ein Redressement, wird die Formel (9.2) zu einer strikten Gleichung und man hat

$$(9.3) \quad \hat{P}(K|\cdot) = [\bar{h}(\xi)/\bar{h}(K)] * \hat{P}(K|h) \quad .^{25}$$

M.a.W., die in Tabelle 7.1 notierten und dort anschließend unter dem hier noch anzufügenden Aspekt der Richtigkeit diskutierten Gewichtungseffekte sind in Anbetracht von Formel (9.3) genau vorhersagbar, da der Quotient der mittleren reduzierten Haushaltsgrößen für alle drei Kategorien bekannt ist ($\bar{h}(\xi) = 1.99!$).

Um dem Leser die Möglichkeit zu geben, zumindest für eine demographische Variable alle für die Gewichtungseffekte in (9.2) verantwortlichen Faktoren im Zusammenhang zu sehen, sei Tabelle 9.1 entsprechend vervollständigt. Die Ergebnisse dieser Vorgehensweise finden sich in der nun folgenden Tabelle 9.4.

Tabelle 9.4: Gewichtungseffekte bei der Variablen "reduzierte Haushaltsgröße"

Kategorie	1	2	>2
$\bar{h}(\xi)/\bar{h}(K)$	1.989	0.9945	0.5853
$\bar{\beta}(\xi)/\bar{\beta}(K)$ - GETAS	0.8438	1.0385	1.2000
$\bar{\beta}(\xi)/\bar{\beta}(K)$ - ZUMA	1.0000	0.9804	1.0000
Korrekturfaktor - GETAS	0.9505	0.9505	0.9630
Korrekturfaktor - ZUMA	1.0050	1.0149	1.0110

Tabelle 9.4 zeigt, daß bei Variablen, die eine beträchtliche Assoziation zu denjenigen Variablen besitzen, welche die Partitionen für ein Redressement bestimmen, keineswegs der Einfluß des Faktors $\bar{\beta}(\xi)/\bar{\beta}(K)$ bestimmend sein muß. Obwohl nämlich bei der GETAS-Gewichtung in den beiden Randkategorien absolut genommen dieser Faktor nicht vernachlässigbar ist (er ist immerhin für rund 15% bzw. 10% "Abweichung" verantwortlich) überwiegt dennoch stark der Einfluß von $\bar{h}(\xi)/\bar{h}(K)$.

Diese wenigen Einzelbeispiele mögen genügen, um zu zeigen, daß es zur Beantwortung der Frage, "ob sich denn überhaupt die gewichtet berechneten Zahlen von den ungewichtet berechneten unterscheiden", keine allgemeine Regel geben kann. Insbesondere, wenn sich diese Frage stellt im Zusammenhang mit komplexen multivariaten Analysen, die sich auf mehrdimensionale Kreuztabellen stützen, verbietet es Formel (9.2) geradezu, pauschal zu sagen, daß sich bestimmte Effekte beim Gewichten einstellen werden oder nicht. Als stets auf den Einzelfall bezogene Grundregel läßt sich jedoch aus (9.2) ablesen, daß die Gewichtungseffekte um so geringer ausfallen, je weniger die zu analysierenden Partitionen bzw. deren konstituierenden Variablen assoziiert sind zur Variablen "reduzierte Haushaltsgröße" bzw. zur jeweils verwendeten Redressement-Partition.

Umgekehrt läßt sich sagen, daß die Gewichtungseffekte bei Gewichten des Typs $w = h * \beta$ um so größer werden, je "näher" die betrachteten Partitionen den das Gewicht bestimmenden Variablen kommen und je mehr dabei eine der Gewichtungskomponenten überwiegt.

Bevor man aus solchen Konsequenzen der Formel (9.2) Handlungsanweisungen zur Benutzung von Gewichten ableitet, muß selbstverständlich

geklärt werden, ob eine Gewichtung die Anteilsschätzungen überhaupt verbessern kann.

Eines der Resultate der vorangegangenen Abschnitte war, daß in bezug auf die "ideale" Stichprobe, auf der Modellseite also, die Gewichtung mit der reduzierten Haushaltsgröße zu einem in unserem Sinne von Repräsentativität brauchbaren Schätzverfahren \hat{Y} (vgl. (7.1)) führt. Es zeigte sich jedoch, daß die bloße Übertragung dieser Gewichtung auf die durch Ausfälle beschädigte Stichprobe zu sehr unbefriedigenden Schätzergebnissen führen kann, die ihrerseits Anlaß sind, Redressementfaktoren einzuführen. Um, wie angekündigt, solchermaßen "aufgebesserte" Schätzverfahren zu vergleichen mit $\hat{Y} = \hat{P}(K^0|h)$, sei an Terminologie eingeführt:

$\xi^0(L)$ = Menge der Personen aus ξ^0 , die nicht in der realisierten Stichprobe ξ enthalten sind (LOSSES);

$n^0(L)$ = (unbekannte) Anzahl der Elemente der Menge $\xi^0(L)$;

$K^0(L)$ = Menge derjenigen Personen aus der Kohorte K , die in ξ^0 , jedoch nicht in ξ enthalten sind;

$k^0(L)$ = (unbekannte) Anzahl der Elemente der Menge $K^0(L)$;

$\bar{h}(\xi^0(L))$ = mittlere reduzierte Haushaltsgröße in $\xi^0(L)$;

$\bar{h}(K^0(L))$ = mittlere reduzierte Haushaltsgröße in $K^0(L)$.

Mit diesen Bezeichnungen gilt - wiederum näherungsweise -

$$(9.4) \quad \hat{P}(K^0|h) \approx \frac{1 + \frac{k^0(L)}{k} * \frac{\bar{h}(K^0(L))}{\bar{h}(K)}}{1 + \frac{n^0(L)}{n} * \frac{\bar{h}(\xi^0(L))}{\bar{h}(\xi)}} * \frac{\bar{\beta}(\xi)}{\bar{\beta}(K)} * \exp\{\dots\} * \hat{P}(K|h\beta),$$

wobei der Korrekturfaktor $\exp\{\dots\}$ gleich demjenigen in Formel (9.2) ist.²⁶

Verglichen mit der Interpretation von Formel (9.2) ist hier allein schon deswegen die Interpretation schwieriger, weil durch die Frage nach der Richtigkeit von $\hat{P}(K|h\beta)$ das - in der Regel unbekannte - Ausfallgeschehen sowohl in der Gesamtstichprobe als auch in der Kohorte explizit einbezogen werden muß. Der dies besorgende erste Faktor auf der rechten Seite der Darstellung zeichnet sich dadurch aus, daß die unbekannten

und nicht notwendig voneinander abhängigen Quotienten $k^o(L)/k$ und $\bar{h}(K^o(L))/\bar{h}(K)$ sich gegenseitig je nach Kohorte in ihrer Wirkung verstärken oder aber neutralisieren können. Eine Prognose zur Richtigkeit von $\hat{P}(K|h\beta)$ erscheint allein von daher schon recht aussichtslos. Dennoch wird sich zeigen, daß praktisch verwertbare Schlußfolgerungen mit Hilfe von (9.4) möglich sind.

In (9.4) sind unbekannt die Größen $\hat{P}(K^o|h)$, $k^o(L)$, $\bar{h}(K^o(L))$, $n^o(L)$ und $\bar{h}(\xi^o(L))$. In jedem Fall schätzbar ist der Nenner des ersten Faktors, indem man Daten des Mikrozensus heranzieht. Auf die Schätzmethode soll hier nicht näher eingegangen werden.²⁷ Wird K durch soziodemographische Merkmale beschrieben, läßt sich auch $\hat{P}(K^o|h)$ zuweilen durch die entsprechenden Prozentzahlen aus dem Mikrozensus - oder aus einer anderen amtlichen Statistik - ersetzen. In dieser günstigen Situation ist dann

$(k^o(L)/k) * (\bar{h}(K^o(L))/\bar{h}(K))$ berechenbar. Man erhält also eine Schätzung über die Ausfallverhältnisse in der Kohorte K .

Leider kann dieser Gedanke hier nicht weiterverfolgt werden, da wegen der Zufallsfehler, die den in alle Rechnungen eingehenden Größen anhaften, erst durch Replikation (auch in bezug auf die Mikrozensus-Daten) von Untersuchungen des Typs ALLBUS 1980 die Größenordnungen von $(k^o(L)/k) * (\bar{h}(K^o(L))/\bar{h}(K))$ mit größerer Sicherheit für ausgewählte Kohorten feststellbar sind.

Lediglich anhand eines durch starke Effekte auffallenden Beispiels sei gezeigt, daß die Ausfallverhältnisse in einer bestimmten Kohorte ganz anders aussehen können als in der Stichprobe ξ :

Für die Kohorte der mindestens 70jährigen Frauen in der Grundgesamtheit entnimmt man der Tabelle 9.2, daß sich mit dem GETAS-Gewicht ein Prozentsatz von 7.5% ergibt. Der Mikrozensus 1980 gibt 8.9% an. $\bar{\beta}(K)/\bar{\beta}(\xi)$ und der Korrekturfaktor $\exp\{\dots\}$ haben die Werte 0.6429 bzw. 0.9671 (vgl. S.171). Setzt man in (9.4) ein, ergibt sich, daß $(k^o(L)/(k) * (\bar{h}(K^o(L))/\bar{h}(K))) \sim 1.6$. Da man nach Tabelle 7.2 annehmen darf, daß sich $\bar{h}(K^o(L))/\bar{h}(K)$ nur unwesentlich von Eins unterscheidet, ist schließlich $k^o(L) \sim 1.6 * k$ oder aber $k/k^o \sim 0.38$, d.h. die "Ausschöpfung" in dieser Kohorte wird mit rund 38% geschätzt. Dies unterscheidet sich deutlich von der insgesamt erreichten "Ausschöpfung" von knapp unter 70%.

Da in den meisten Fällen für die interessierenden Kohorten der Wert $\hat{P}(K^0|h)$ unbekannt ist, wird eine gute Schätzung für die Ausfallverhältnisse in einer bestimmten Kohorte eher die Ausnahme sein. Man kann dann nicht mit Bestimmtheit sagen, daß eine Gewichtung mit $w = h * \beta$ die Schätzwerte verbessert oder verschlechtert. Insbesondere ist dies nicht möglich in bezug auf die Frage, ob ein Redressement verwendet werden soll oder nicht. Jedoch lassen sich ähnlich wie beim Vergleich "ungewichtet - gewichtet" Kriterien angeben, die für den Einsatz einer Gewichtung als Entscheidungshilfe herangezogen werden können.

Für die Qualität des Schätzers $\hat{P}(K|h\beta)$ ist entscheidend, wie weit das Produkt der ersten drei Faktoren auf der rechten Seite von (9.4), der "Abstandsfaktor", von Eins abweicht. Verschlechternd wirkt es deshalb in jedem Fall, wenn Redressementeffekt und Ausfalleffekt beide entweder kleiner oder größer als Eins sind. Verbessernd kann es wirken, wenn sich Ausfalleffekt und Redressementeffekt gegenseitig aufheben, was ja für jede Kohorte aus der Redressement-Partition genau zutrifft. Zwei Beispiele mögen diese beiden Feststellungen erläutern.

Für die Kohorte derjenigen Personen aus der Grundgesamtheit, die in Haushalten der reduzierten Größe von mindestens Drei leben, ist nach Tabelle 9.4 für das GETAS-Gewicht $\bar{\beta}(\xi)/\bar{\beta}(K) = 1.2000$ und der Korrekturfaktor hat die Größe 0.9630; das Produkt ergibt sich also zu rund 1.16. Die nicht unplausible Annahme, daß in dieser Kohorte die reduzierten Haushaltsgrößen bei den nicht-stichprobenneutralen Ausfällen im Schnitt höher liegen als bei den realisierten Interviews - und dies in der Gesamtstichprobe gerade umgekehrt ist - führt dazu, daß auch der Faktor für den Ausfalleffekt größer als Eins sein muß. Das Redressement verstärkt also hier den Ausfalleffekt. Dies zeigt sich auch sehr schön am Vergleich der Prozentzahlen, die zum einen gewichtet mit der reduzierten Haushaltsgröße allein und zum anderen gewichtet mit dem GETAS-Gewicht berechnet wurden, vgl. Tabelle 9.1: dem prozentualen Anteil von 29.9% liegt in bezug auf den "wahren Wert" derselbe Faktor für den Ausfalleffekt zugrunde, wie dem mit dem GETAS-Gewicht berechneten Anteil von 25.8%. Man sieht dies unmittelbar an Formel (9.4), wenn man β_i identisch Eins setzt; $\hat{P}(K|h\beta)$ wird dann zu $\hat{P}(K|h)$ oder m.a.W., das Produkt der drei letzten Faktoren auf der rechten Seite von (9.4) ist ungefähr gleich $\hat{P}(K|h)$. Die Hinzufügung des Redressementfaktors bedeutet also,

daß $\hat{P}(K|h)$ "nach unten" verschoben wird und die Division von $\hat{P}(K|h)$ durch 1.16 liefert in der Tat den Wert von 25.8.

Gerade umgekehrt wirkt das Redressement des GETAS-Gewichts bei der Kohorte derjenigen Personen der Grundgesamtheit, die in Haushalten der reduzierten Größe Eins leben. Hier ist unmittelbar einsichtig, daß der Ausfalleffekt-Faktor größer als Eins sein muß, was ausgeglichen wird durch $\bar{\beta}(\xi)/\bar{\beta}(K) = 0.8438$ (vgl. Tabelle 9.4). Dies führt in der Tat zu dem verbesserten Schätzwert von 16.2%, verglichen mit den 12.9%, die sich bei Gewichtung mit der reduzierten Haushaltsgröße allein ergeben.

Beide Beispiele zusammen zeigen, daß ein Redressement selbst "innerhalb" einer Variablen sowohl verbessernd als auch verschlechternd wirken kann, und zwar in bezug auf die Gewichtung mit der reduzierten Haushaltsgröße allein. Ein Redressement, obwohl es ja für Ausfälle ausgleichen soll, kann sogar fast wirkungslos gegen Fehlschätzungen bleiben, wenn in einer Kohorte die Ausfallverhältnisse sich von denen in der Stichprobe sehr stark unterscheiden. Dies gilt z.B. für die Kohorte der Personen in Haushalten der reduzierten Größe Zwei. Darin ist für das GETAS-Gewicht $\bar{\beta}(\xi)/\bar{\beta}(K) = 1.0385$ (vgl. Tabelle 9.4), jedoch der Ausfalleffekt-Faktor ist wegen der geringen Ausfälle in dieser Kohorte so erheblich kleiner als Eins, daß eine deutliche Überschätzung des Mikrozensus-Wertes resultiert.

Insgesamt zeigen diese Beispiele für einen Variablentyp aus dem demographischen Bereich, daß

- nicht erwartet werden kann, daß die definierten Redressementfaktoren auch in Partitionen, die von der Redressement-Partition verschieden sind, zu einer gleichmäßigen Verbesserung der Schätzergebnisse führen, selbst wenn eine nicht vernachlässigbare Assoziation zwischen beiden Partitionen unterstellt werden muß;
- die Redressement-Wirkung in einzelnen Fällen von den disproportionierenden Wirkungen nicht-stichprobenneutraler Ausfälle zunichte gemacht werden kann.

Praktisch kann man diese Ergebnisse so berücksichtigen, daß man für die Elemente einer Partition, also Element gleich Kohorte, die Größen $\bar{h}(K)$ und $\bar{\beta}(K)$ kontrolliert. Liegt $\bar{h}(K)$ für alle K nahe bei $\bar{h}(\xi)$ und schwankt $\bar{\beta}(K)$ nur unwesentlich um $\bar{\beta}(\xi)$, wird man in der Mehrzahl der Fälle mit

"guten" Schätzwerten $\hat{P}(K|h\beta)$ rechnen dürfen. Konzentrieren sich jedoch kleine bzw. große Werte von $\hat{h}(K)$ auf einzelne Kohorten und ist dies mit größeren Bandbreiten des Wertebereichs der $\hat{\beta}(K)$ verbunden, muß damit gerechnet werden, daß trotz der Gewichtung deutliche Fehlschätzungen auftreten können.

Diese Schlußweisen gelten selbstverständlich auch für Variablen oder Partitionen, die mehr dem "inhaltlichen" Bereich zugeordnet werden. Erfahrungsgemäß - nach den Formeln (9.2) und (9.4) sogar hochplausiblerweise - sind dort die Gewichtungseffekte im Schnitt geringer als bei Partitionen, die eher der Demographie zuzuordnen sind; vgl. auch Tabelle 9.3. Wie man Tabelle 9.5 entnimmt, können die Schwankungen von $\hat{h}(K)$ und $\hat{\beta}(K)$ sogar ausgesprochen gering sein. Jedoch sollte man sich stets vergegenwärtigen, daß dies nicht in jedem Fall bedeuten muß, daß der Ausfalleffekt-Faktor aus (9.4) ebenfalls sehr nahe bei Eins liegt, daß also $\hat{P}(K|h\beta)$ durchaus noch verzerrt sein kann.

Tabelle 9.5: "In Behörden wird man wie eine Nummer behandelt"

Kategorie	1	2	3	4	5	6	7
$\hat{h}(\xi)/\hat{h}(K)$	0.9950	0.9830	1.0102	1.0153	0.9900	0.9900	0.9950
$\hat{\beta}(\xi)/\hat{\beta}(K)$ (GETAS)	0.9643	1.0000	1.0000	1.0385	0.9818	1.0385	1.0000
$\hat{\beta}(\xi)/\hat{\beta}(K)$ (ZUMA)	1.0204	1.0000	0.9804	1.0000	0.9434	1.0417	1.0000
Korrekturfaktor (GETAS)	1.0118	0.9914	0.9904	0.9781	0.9815	0.9925	1.0235
Korrekturfaktor (ZUMA)	0.9948	1.0100	0.9997	1.0258	1.0006	1.0101	1.0050

10. Die Ergebnisse der Abschnitte 7 bis 9 im Überblick

In den Abschnitten 4 und 5 wurde ein mathematisches Modell für den Stichprobenplan des ALLBUS 1980 dazu benutzt, ein Schätzverfahren für Mittelwerte - nämlich $\hat{\bar{Y}}$ in (4.4) - mit Hilfe von Auswahlwahrscheinlichkeiten zu begründen und in bezug auf seine Qualität zu untersuchen.

Gegenstand des Abschnittes 7 war es aufzuzeigen, daß die idealisierenden Annahmen dieses mathematischen Ziehungsmodells in der Praxis insbesondere wegen des Ausfallgeschehens in der Regel nicht mehr zutreffen. Daß also auch die Anteilsschätzung \hat{Y}^a nicht ohne weiteres angewandt werden kann auf die Daten der realisierten Stichprobe.

Die Realitätsferne des mathematischen Modells macht dieses jedoch keineswegs bedeutungslos. Mangels angemessener Modelle, die das Ausfallgeschehen mit einschließen, ist es vielmehr naheliegend zu versuchen, den "Abstand" zwischen den Werten von \hat{Y}^a aus dem Modell und denjenigen Schätzwerten zu bestimmen, die aus praktisch weithin verwendeten Schätzverfahren resultieren. Da die Qualität von \hat{Y}^a im Prinzip beschreibbar ist, ergibt sich so die Möglichkeit zu beurteilen, ob die zwei im Abschnitt 8 angegebenen Anteilsschätzer - die durch ein ganz bestimmtes Verfahren der Schätzverbesserung (Redressement) aus \hat{Y}^a gewonnen wurden - die erhoffte Güte haben.

Die beiden Modifikationen von \hat{Y}^a bestehen aus einer Gewichtung des Typs $w = h * \beta$ - an Stelle einer Gewichtung mit h allein; dabei ist h die reduzierte Haushaltsgröße (vgl. Anmerkung 11) und im Faktor β wird das jeweils verwendete Redressement zusammengefaßt.

Gegenstand des Abschnittes 9 war es,

- die numerische Wirkung dieser Gewichtung im Vergleich zu ungewichteter Berechnung in systematischer Weise darzustellen (vgl. Formel (9.2)) und
- den "Abstand" zwischen solchermaßen gewichtet berechneten Anteilswerten und den Werten von \hat{Y}^a anzugeben (vgl. Formel (9.4)).

Für den rein numerischen Effekt der Gewichtung (also "gewichtet" vs. "ungewichtet") konnte aus (9.2) der Schluß gezogen werden, daß es nicht möglich ist, im vornhinein pauschal zu sagen, daß sich ganz bestimmte Gewichtungseffekte einstellen werden; die Einflüsse der Variablen h und β können sich nämlich in unterschiedlicher und kaum vorhersehbarer Weise überlagern, d.h. sich gegenseitig verstärken oder aber neutralisieren. Als allgemeine Regel ergab sich jedoch, daß

"die Gewichtungseffekte um so geringer ausfallen, je weniger die zu analysierenden Partitionen bzw. deren konstituierenden Variablen assoziiert

sind zur Variablen 'reduzierte Haushaltsgröße' bzw. zur jeweils verwendeten Redressement-Partition."

Der "Abstand" zwischen \hat{Y} und den vermöge der Gewichtung $w = h * \beta$ berechneten Schätzwerten wird nach Formel (9.4) in verwickelter Weise durch das Zusammenspiel von Redressement-Effekten und Ausfalleffekten bestimmt. Wie beim Vergleich "ungewichtet" vs. "gewichtet" können durch Überlagerung gegenseitige Verstärkung aber auch Neutralisierung eintreten. Für die praktische Anwendung der Gewichtung(en) sind auch hier also pauschale Aussagen über die Güte der gewichtet berechneten Schätzungen nicht möglich. Auf den Einzelfall bezogen läßt sich jedoch sagen, daß bei einer Gewichtung mit w das Verhalten der Mittelwerte der Variablen h und β , berechnet für die interessierenden Kohorten, von zentraler Bedeutung ist. Liegen diese Werte durchweg nahe bei den entsprechenden Mittelwerten über die Gesamtstichprobe, wird man in der Mehrzahl der Fälle "gute", d.h. nahe bei \hat{Y} gelegene Schätzungen unterstellen können. Konzentrieren sich dagegen kleine bzw. große Kohortenmittelwerte von h auf einzelne Kohorten und fällt dies zusammen mit größeren Wertebereichen der zugehörigen Kohortenmittelwerte von β , muß trotz Gewichtung mit deutlichen Fehlschätzungen gerechnet werden.

Anmerkungen

- 1 Als Privathaushalt - in der Begrifflichkeit der amtlichen Statistik - gilt jede Gemeinschaft von Personen, die zusammen wohnt und gemeinsam wirtschaftet. Die Personen innerhalb eines Privathaushalts sind also nicht notwendig miteinander verwandt.
- 2 Von allen mindestens 18jährigen Personen mit deutscher Staatsangehörigkeit in der Bundesrepublik und West-Berlin leben rund 1.5% in Anstalten.
- 3 ADM = Arbeitskreis Deutscher Marktforschungsinstitute; ein Verzeichnis seiner Mitglieder findet sich in Schaefer (1979: 131ff.); vgl. auch Abschnitt 3.2.1.
- 4 Es werden in diesem Abschnitt der Übersichtlichkeit halber Vereinfachungen vorgenommen; so wurden z.B. vor der Ziehung aus der Datei der Stimmbezirke zu kleine Einheiten eliminiert, indem sie größeren zugeschlagen wurden. Im Detail werden die Prozeduren bei der Stichprobenziehung in den Abschnitten 3.2.1 bis 3.2.3 dargestellt.
- 5 Zum Begriff der Wohnbevölkerung vgl. Abschnitt 2.

- 6 So ist z.B. zu berücksichtigen gewesen, daß die Bedeutungsgewichte s einiger Primäreinheiten die Schrittweite überstiegen. Diese Einheiten waren also mit Sicherheit in der Stichprobe vertreten, die größeren sogar mehrfach. Da sie selbstverständlich nur einfach "zählen", war vor der Stichprobenziehung ein "Brutto" von mehr als 25 200 festzulegen, um auf eine effektive Stichprobengröße von 25 200 zu kommen.
- 7 Von diesen 630 Primäreinheiten entstanden 73 durch Synthetisierung.
- 8 Die Befragungen zum ALLBUS 1980 fanden in 426 verschiedenen Gemeinden statt und von diesen hatten 40, also 9.4%, weniger als 2 000 Einwohner. Man beachte jedoch, daß zum Gebietsstand 1.1.1977 rund 65% der Gemeinden in der Bundesrepublik weniger als 2 000 Einwohner hatten. Diese beiden sehr unterschiedlichen Relationen erklären sich durch die Ziehungsweise der Primäreinheiten (vgl. Abschnitt 3.2.1): Die Wahrscheinlichkeit dafür, daß eine Gemeinde vermöge einer in ihr liegenden Primäreinheit "getroffen" wird, ist proportional zur Summe der Bedeutungsgewichte der in dieser Gemeinde enthaltenen Primäreinheiten. Bedenkt man nun, daß 1977 rund 8% der Wohnbevölkerung in Gemeinden mit weniger als 2 000 Einwohnern lebte, so wird die Relation 426/40 schließlich hochplausibel.
- 9 Ist eine Grundgesamtheit G in k sich nicht überschneidende Teilgruppen (Kohorten, Untergruppen) C_1, \dots, C_k , aufgeteilt, so nennt man das System der C_1, \dots, C_k eine Partition der Grundgesamtheit G mit den Elementen $C_i, i=1, \dots, k$.
Die Grundauszählung einer dichotomen Variablen V z.B. generiert in natürlicher Weise eine Partition:
 C_1 = Menge der Personen, für die V den Wert Null hat,
 C_2 = Menge der Personen, für die V den Wert Eins hat.
- Liegt, um ein weiteres Beispiel zu nennen, eine mehrdimensionale Kreuztabelle vor, die etwa dem Kreuzprodukt der Variablen V_1, V_2, V_3 und V_4 entspricht, und hat die Variable V_j $m(j)$ Ausprägungen, $j=1,2,3,4$, so ergibt sich eine Partition mit $m(1)*m(2)*m(3)*m(4)$ -Elementen; jedes dieser Elemente ist dadurch definiert, daß es alle diejenigen Personen umfaßt, die eine ganz bestimmte Merkmalskombination der vier Variablen aufweisen.
- 10 Der Faktor N/S läßt sich als durchschnittliche Haushaltsgröße in bezug auf die Grundgesamtheit interpretieren, also $N/S \sim 2$.
- 11 Hier und im folgenden soll "reduzierte Haushaltsgröße" die Anzahl der Personen aus der Grundgesamtheit im Haushalt sein.
- 12 Im Jahre 1981 erwarb ZUMA vom Statistischen Bundesamt Tabellen aus dem Mikrozensus 1980, die auf die beim ALLBUS 1980 relevanten Grundgesamtheiten sowohl auf Haushalts- als auch auf Personenebene abgestimmt waren.
- 13 In diesem Zusammenhang sei daran erinnert, daß im Beispiel zum Größenproportionalen Ziehen im Abschnitt 3.2.1 die Varianz von \bar{Y} gerade mit Hilfe der Gleichung (5.1) aufgrund der bekannten fünf Werte $P(\xi)$ berechnet wurde.
- 14 Da sicher in keiner Primäreinheit mehr als fünfmal so viele Zielpersonen wie Pivathaushalte vorhanden sind, also stets $H_i \leq 5 * s_i$ gilt, variiert Y_i/s_i zwischen Null und höchstens 5.

- 15 Alle im folgenden verwendeten populationsbezogenen Zahlen beziehen sich sowohl auf Haushalts- als auch auf Personenebene auf diese Sonderauszählung, vgl. auch Anmerkung 2.
- 16 Daß beide Annahmen zusammen mit der Technik der Abschätzung nach oben vermutlich auf eine noch zu große Schranke führen, kann hier nicht weiter diskutiert werden.
- 17 Die Annahme, daß $Y_{ij} = 0$ oder $Y_{ij} = 1$, ist selbstverständlich erst für $h_{ij} = 4$ notwendig.
- 18 Darunter sei hier verstanden $\{E(\bar{Y} - \bar{Y})^2\}^{1/2} * 100/p$.
- 19 Man beachte hier, daß $Y_i/s_i = (Y_i/H_i) * (H_i/s_i) = p * D_i$.
- 20 Es sei daran erinnert, daß lediglich der weniger schwerfälligen Darstellung wegen die Diskussion sich hier und auch im folgenden stets auf die Schätzung von Anteilen bezieht.
- 21 Man beachte, daß in diesem Fall die Variable mit sich selbst gewichtet wird.
- 22 Hier und im folgenden gibt das Subskript unter dem Summenzeichen an, über welchen Bereich der Stichprobe sich die Summation erstreckt.
- 23 Genauer müßte unter dem Summenzeichen im Zähler der Schnitt der beiden Mengen K und ξ stehen.
- 24 Der Korrekturfaktor $\exp\{\dots\}$ liegt in den allermeisten Fällen so nahe bei Eins, daß die Sprechweise "in aller Regel" berechtigt ist. Im übrigen gilt dies, solange nur $|r(\xi, h, \beta)|$ und $|r(K, h, \beta)|$ hinreichend klein sind, was bei den üblichen Redressements und für "vernünftige" Kohorten K stets unterstellt werden kann.
- 25 Dies ist eine triviale(!) Identität, wie man durch Nachrechnen sofort feststellt.
- 26 Das Produkt der ersten drei Faktoren der rechten Seite von (9.4) stellt den angekündigten "Abstand" zwischen $\bar{Y} = \hat{P}(K^o|h)$ und $\hat{P}(K|h\beta)$ dar.
- 27 Als Schätzung für $n^o(L)/n$ ergibt sich der Wert 0.470; für $\bar{h}(\xi^o(L))$ erhält man 1.64. Da $h(\xi) = 1.99$ (Mikrozensus 1980: 1.88) hat der Nenner den geschätzten Wert $1 + 0.47 * (1.64/1.99) = 1.39$.

Literatur

- Anderson, Harald, 1979: On nonresponse bias and response probabilities. Scandinavian Journal of Statistics, 6: 107-112.
- Boustedt, Olav, 1975: Stadtreionen in der Bundesrepublik Deutschland 1970. Hannover: Akademie für Raumforschung und Landesplanung, Band 103.
- GETAS, 1980: NATIONALER SOZIALER SURVEY 1979/1980 - Zur technischen Durchführung -. GETAS, Gesellschaft für angewandte Sozialpsychologie, Bremen.

- Hansen, Morris, H. und William N. Hurwitz, 1943: On the theory of sampling from finite populations. *Annals of Mathematical Statistics*, 4: 333-362.
- Hoffmeyer-Zlotnik, Jürgen, 1981: Zur Konstruktion eines neuen Stadtindex. *ZUMACHRICHTE*, 9: 47-52.
- Horvitz, D., G. und D.J. Thompson, 1952: A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47: 663-685.
- Kish, Leslie, 1949: A procedure for objective respondent selection within the household. *Journal of the American Statistical Association*, 44: 380-387.
- Kish, Leslie, 1965: *Survey sampling*. New York: Wiley.
- Konijn, H.S., 1973: *Statistical theory of sample survey design and analysis*. Amsterdam: North-Holland.
- Neubäumer, Renate, 1982: *Die Eigenschaften verschiedener Stichprobenverfahren bei wirtschafts- und sozialwissenschaftlichen Untersuchungen*. Frankfurt a.M.: Lang.
- Raj, Des, 1968: *Sampling theory*. New York: McGraw-Hill.
- Rao, J.N.K.; Scott, A.J., 1981: The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness of fit and independence in two-way tables. *Journal of the American Statistical Association*, 76: 221-230.
- Saerndal, Carl-Eric, 1978: Design-based and model-based inference in survey sampling. *Scandinavian Journal of Statistics*, 5: 27-52.
- Schaefer, Felix, 1979: *Muster-Stichproben-Pläne*. München: Moderne Industrie.
- Statistisches Bundesamt, 1979: *Bevölkerung und Erwerbstätigkeit, Reihe 3, Haushalte und Familien*: 15, Mainz: Kohlhammer.
- Sudman, Seymour, 1976: *Applied sampling*. New York: Academic Press.
- ZUMA, 1979: *Ausschreibung zum Nationalen Sozialen Survey 1980*. Zentrum für Umfragen, Methoden und Analysen, Mannheim.

ERFASSEN VON WOHNQUARTIERSVARIABLEN - EIN MITTEL ZUR SOZIOLOGISCHEN ZUORDNUNG DER WOHNBEVÖLKERUNG

Jürgen H.P. Hoffmeyer-Zlotnik

1. Einleitung

Die 'Allgemeine Bevölkerungsumfrage der Sozialwissenschaften 1980' (ALLBUS 1980) bot eine Möglichkeit, Hintergrundmerkmale über einen Befragten nicht durch die Antwort des Befragten selbst, sondern auch über ein Beobachtungsprotokoll zu erfassen. Dieses Beobachtungsprotokoll war ein dem Befragungsinstrument angehängtes, vom Interviewer auszufüllendes Instrument zur Erfassung von Wohnquartiersvariablen.

Nun befinden sich in den Demographieteilen sehr vieler Umfragen Fragen zur Wohnquartiersbeschreibung. Auch die ZUMA-Standarddemographie (vgl. Pappi 1979) enthält Fragen zum Wohngebiet und zum Wohngebäude des Zielhaushaltes. Genutzt werden diese Informationen bisher als Hintergrundmerkmale bzw. zur Milieuschilderung. Und genau bei dieser Milieuschilderung klingt die Annahme durch: "Sag mir, wo Du wohnst und ich sag Dir, wer Du bist". Träfe diese Annahme zu, dann müßte es auch möglich sein, über die Beschreibung eines Wohngebietes Rückschlüsse auf die soziologischen Merkmale der Bewohner dieses Gebietes zu wagen.

Bisher wurden die Variablen einer Wohnquartiersbeschreibung in der Regel über die subjektive Einschätzung der Befragten erhoben. Sie lassen sich aber auch in Form eines vom Interviewer auszufüllenden Beobachtungsprotokolles erfassen. Diese separate Aufzeichnung der Wohnquartiersbeschreibung in einem durch den Interviewer ausgefüllten Beobachtungsprotokoll würde dem Forscher eine weniger durch subjektive Erfahrungen und Vorurteile verzerrte Schilderung des Milieus liefern, in dem der zu Befragende wohnt und lebt. Diese Wohnquartiersbeschreibung ließe dann Rückschlußmöglichkeiten auf den Zielhaushalt des zu Befragenden zu. Da der Interviewer zum Ausfüllen des Beobachtungsprotokolls nicht die Hilfe des zu Befragenden braucht, sind solche Rückschlußmöglichkeiten auch dann gegeben, wenn eine Befragung nicht zustande kommt.

Wenn aber Wohnquartiersbeschreibungen Rückschlüsse auf die in einem Wohnquartier Lebenden zulassen, dann sollte es auch möglich sein, über eine Wohnquartiersbeschreibung jene Personengruppen, die sich hinter den Befragungsausfällen verbergen, einzukreisen. Allerdings wird diese Methode des Rückschlusses auch nur eine Annäherung an die soziologischen Merkmale des Nicht-Befragten bringen: Zu identifizieren ist nur die in einem Wohnquartier dominante Gruppe (und der Befragte/zu Befragende muß natürlich nicht unbedingt der dominanten Bevölkerungsgruppe angehören). Der Rückschluß auf die soziologischen Merkmale eines Quartiersbewohners über eine Quartiersbeschreibung ist aber wesentlich aussagekräftiger als es zu einem Idealwert hin verzerrende Hochrechnungen über den Vergleich mit der amtlichen Statistik oder als es gar Vermutungen sein könnten. Eine absolut sichere Methode ist der Rückschluß auf die Quartiersbewohner über eine Quartiersbeschreibung also nicht.

Die einzige, annähernd sichere Methode, soziologische Merkmale einer Person zu erheben, wäre die Einordnung des bei der Befragung Ausgefallenen über anderweitig von diesem selbst zur eigenen Person gemachte Aussagen. So wäre theoretisch eine Einordnung mittels Zugriff auf personenbezogene Daten der Einwohnermeldedatei denkbar. Dies ist jedoch eine rein hypothetische Möglichkeit, da dem nicht nur Datenschutzregelungen entgegenstehen, sondern auch dies nur bei kleinen oder regional eng begrenzten Stichproben technisch durchführbar wäre. Als ein praktikables Verfahren zur spezifischen Zuordnung von Ausfällen bleibt daher nur die Einschätzung der nicht befragbaren Zielperson über vom Interviewer zu erhebende, objektiv beschreibbare Merkmale.

Im folgenden wird nun dargestellt, wie Wohnquartiersvariablen als Mittel zur Bestimmung soziologischer Merkmale herangezogen werden können und wie sie sich in einem ersten Test im ALLBUS 1980 bewährt haben. Selbstverständlich mußte ein erster Versuch, Interviewer auch als Beobachter einzusetzen, ein Experiment bleiben, welches dem Forscher die Richtung zu weisen hatte, in die die Weiterentwicklung zu gehen habe. Daher wird hier nicht nur über die Erhebung von Wohnquartiersvariablen im ALLBUS 1980 berichtet, sondern es wird auch der Folgeversuch, der zweite Einsatz eines Beobachtungsprotokolls zur Wohnquartiersbeschreibung mit einem modifizierten Erhebungsinstrument, dargestellt (Hoffmeyer-Zlotnik

1981a). Dennoch steht am Ende kein erprobtes, sondern nur ein in der Erprobung befindliches Instrument.

2. Sozialökologischer Hintergrund

Ein Rückschließen über die Beschreibung eines städtischen Teilgebietes auf die dort ansässige Wohnbevölkerung ist dann theoretisch sinnvoll, wenn man von der sozialräumlichen Differenzierung von Bevölkerungsgruppen ausgeht. Man setzt voraus, daß sich unterschiedliche Gruppen in verschiedenen Wohnquartieren ansiedeln, weil Wohnquartiere aufgrund ihrer spezifischen Lage innerhalb einer Stadt, der Art ihrer Bebauung und der damit verbundenen Raumansprüche sowie der unterschiedlichen Ausstattung mit sozialen Infrastruktureinrichtungen einen jeweils verschieden hohen Statuswert aufweisen. Bewohner eines Wohnquartiers neigen daher dazu, den Statuswert ihres Gebietes nicht durch den Zuzug von Personengruppen, welche mit einem niedriger eingeschätzten sozialen Status versehen sind, abwerten zu lassen (Duncan und Duncan 1957; Hoffmeyer-Zlotnik 1979). Denn in der Praxis gilt eine "Adresse" solange als eine "gute Adresse", wie die Haushalte, die dort wohnen, den sozialen Status haben, den man der Adresse zuschreibt.

Dieser sozialräumlichen Differenzierung liegen im wesentlichen folgende Annahmen einer sozialökologischen Theorie von E.W. Burgess (1925, 1929, nach Friedrichs 1977: 101) zugrunde:

Expansion (physische Ausdehnung) und Wachstum (der Bevölkerung) führen zu den gegensätzlichen wie komplementären Prozessen der Konzentration und Dezentralisation.

Jede Stadt weist eine innere Gliederung auf, die sich idealtypisch in einem Modell konzentrischer Zonen abbilden läßt (siehe Abbildung 1).

Die Nutzungen und die Bevölkerungsgruppen sind nicht gleichmäßig über das metropolitane Gebiet verteilt, es überwiegen vielmehr in jeder Zone bestimmte Nutzungen und bestimmte Bevölkerungsgruppen.

Die gegebene Differenzierung der Berufe führt zu räumlichen Verschiebungen und Trennungen der einzelnen Berufe (Segregation).

Die Gebiete einer Stadt betonen tendenziell bestimmte Merkmale ("Züge") und ziehen bestimmte Individuen an.

Die Gebiete "entwickeln" die in ihnen lebenden Individuen und differenzieren sich so selbst weiter.

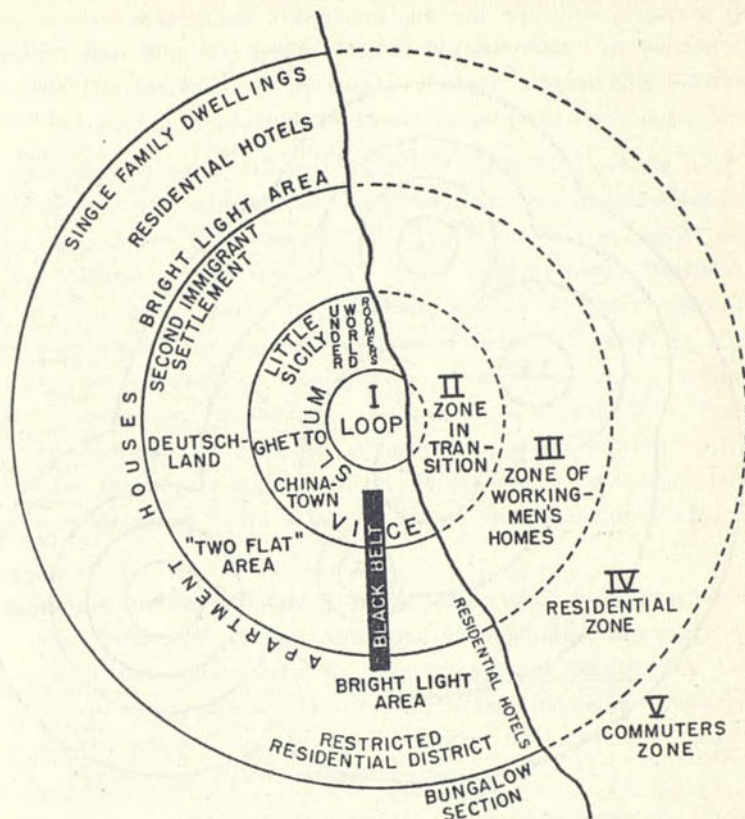
Die gegenwärtige reorganisierte Großstadt ist ein zentralisiertes (weil auf unterschiedliche Zentren mit unterschiedlichen Funktionen, welche in Hierarchie zueinander stehen, ausgerichtetes) dezentrales System lokaler Gemeinden (siehe Abbildung 2).

Die vorgestellten Hypothesen von Burgess wurden innerhalb von sechs Jahrzehnten für marktwirtschaftlich orientierte Industrienationen nicht widerlegt (vgl. Theodorson 1961; Friedrichs 1977; 1978; Hamm 1977; Hoffmeyer-Zlotnik 1977). Sie machen deutlich, daß eine Stadt aus einer Vielzahl unterschiedlicher Teilgebiete besteht (siehe Abbildung 2), die von einer jeweils relativ homogenen Bevölkerungsgruppe bewohnt werden. In einzelnen, kleineren Teilgebieten einer Stadt ist jeweils eine Gruppe, die einer bestimmten sozialen Schicht zugerechnet werden kann, dominierend. Die Palette der jeweils mit einem unterschiedlich bewerteten sozio-ökonomischen Status versehenen städtischen Teilgebiete, von einer deren sozio-ökonomischen Status adäquaten Bewohnergruppe dominiert, reicht von der "gold coast" (früher das großbürgerliche Westend, heute die Villenviertel und Wohngebiete des alten Patriziats und der Reichen, Wohngebiete, die in Mitteleuropa - wegen der vorherrschenden Westwinde - immer im Südwesten einer Großstadt liegen) bis zum "slum" (in der Regel die noch nicht modernisierten bzw. noch nicht sanierten oder ungenutzten, an den zentralen Geschäftsbereich angrenzenden Altbauwohngebiete, die vor dem Ersten Weltkrieg bezugsfertig geworden sind).

Die hier skizzierte räumliche Verteilung städtischer Bevölkerung hinsichtlich ihrer sozialen Differenzierung ist gut untersucht. Dank dieser Untersuchungen der Sozialökologen ist ein großer Teil jener Gesetzmäßigkeiten bekannt, nach denen die Prozesse verlaufen, die die sozialräumlichen Strukturen organisieren, verändern und reorganisieren. Weil diese Gesetzmäßigkeiten der Prozeßabläufe nicht nur für die USA, sondern auch für Mitteleuropa belegt sind (z.B. Friedrichs 1978; Hamm 1977), lassen sich bestimmten städtischen Teilgebieten ganz bestimmte dominante soziale Gruppen zuschreiben und Prognosen über die Veränderung der sozialräumlichen Verteilung verschiedener städtischer Bevölkerungsgruppen stellen.

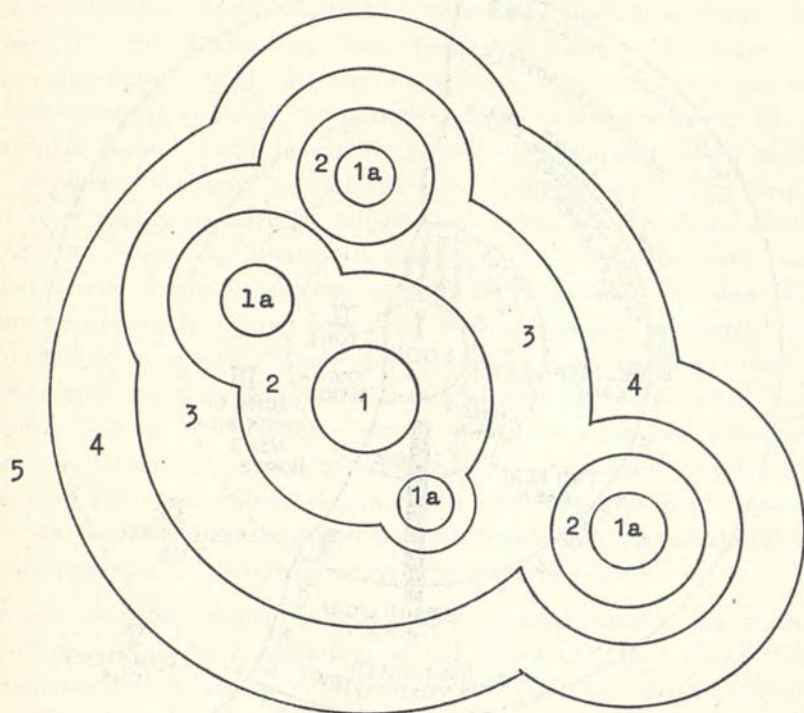
Allerdings muß hier gleich vor einem Fehlschluß gewarnt werden: Nachbarn gehören nicht unbedingt immer der gleichen sozialen Schicht an, weisen nicht unbedingt gleiche soziologische Gruppenmerkmale auf, auch

Abbildung 1: Stadtentwicklungsmodell von E.W. Burgess (1925: 55) am Beispiel der Stadt Chicago/Ill.



- I Loop = zentraler Geschäftsbezirk
- II Zone des Übergangs, City-Erweiterungs-Gebiet
- III Arbeiterwohngebiet, Mietshausdistrikt
- IV Wohngebiet der Mittel- und Oberschichten
- V Region der Vorstädte

Abbildung 2: Stadtentwicklungsmodell von J. Hoffmeyer-Zlotnik (1977: 18)



- 1 = Hauptzentrum
- 1a = Neben-, Unterzentren
- 2 = "zone in transition"
- 3 = Arbeiterwohngebiet, Mietshausdistrikt
- 4 = Wohngebiet der Mittel- und Oberschichten
- 5 = Region der Vorstädte

wenn sie nebeneinander und im gleichen Typ Haus wohnen (vgl. Zapf 1969: 126ff.)! Das heißt, betrachtet man eine zu kleine Untersuchungseinheit, so erweist sich ein auf höherer Aggregatebene relativ homogen erscheinendes städtisches Teilgebiet als ein sehr heterogenes Gebilde; betrachtet man eine zu große Untersuchungseinheit, so werden die bestehenden Unterschiede durch einen zu stark nivellierenden Effekt verwischt: bestehende Gruppenunterschiede fallen nicht mehr auf.

Ziel muß es also sein, Städte hinsichtlich ihrer Wohngebiete so stark zu untergliedern, daß unterschiedliche Wohngebiete mit unterschiedlichen Bevölkerungsgruppen sich voneinander abheben und ein Rückgriff auf sogenannte "natural areas" (vgl. Hatt 1946) möglich wird. "Natural areas" sind in sich geschlossene Wohngebiete gleicher Art der Bebauung, in denen eine Bevölkerungsgruppe die dominierende ist. Wenn es aber möglich ist, unterschiedliche Wohngebiete voneinander durch optische Erscheinungsmerkmale zu trennen, also "natural areas" über ökologische Variablen sichtbar voneinander abzugrenzen, dann müssen auch unterschiedliche Bevölkerungsgruppen mit verschiedenen soziologischen Merkmalen den einzelnen "natural areas" (oder "Wohnquartieren") zuzuordnen sein.

Damit ist die hier anstehende Aufgabe jedoch noch nicht erschöpfend gelöst, denn gefordert ist das Auffinden gleicher Bevölkerungsgruppen in unterschiedlichen Wohnquartieren unterschiedlicher Städte. Das bedeutet, daß jene Variablen, die zur Bestimmung unterschiedlicher "natural areas" herangezogen werden, in allen Städten oder zumindest allen Städten des gleichen Typus, das gleiche messen müssen.

So betrachtet, ergibt sich (bei hoher Treffersicherheit) die Möglichkeit, über eine Gebietsbeschreibung, auf die in einem städtischen Teilgebiet dominante Bevölkerungsgruppe rückzuschließen. Voraussetzung hierfür ist jedoch, daß sich städtische Teilgebiete mit gleichem Status in unterschiedlicher Lage und in unterschiedlichen Städten mittels gleicher Variablen bestimmen lassen.

Mit welchen Variablen lassen sich unterschiedliche städtische Teilgebiete mit unterschiedlichem sozialen Status (dominiert von unterschiedlichen Bevölkerungsgruppen) voneinander abgrenzen?

Es sind Variablen, die

1. die Lage eines städtischen Teilgebietes in der Stadt bzw. zur Stadt beschreiben,
2. die Art und die Qualität der Verkehrerschließung, d.h., die Erreichbarkeit eines städtischen Teilgebietes, angeben,
3. die Bebauungsdichte aufzeigen,
4. das Baualter und den Bauzustand der vorzufindenden Gebäude dokumentieren,
5. die (überwiegende) Art der Landnutzung aufzeigen und beschreiben, ob es sich um ein reines Wohngebiet oder um ein Mischgebiet mit hohem Anteil an a) Handelsflächen, b) Gewerbeflächen, c) Industriefläche, d) Verwaltungsfläche handelt,
6. den Stadttyp beschreiben. Das heißt, um welche Art Stadt es sich handelt: Ob es eine Groß-, Mittel-, Kleinstadt oder Dorf ist (Größe); ob sie im ländlichen Raum oder in einem städtischen Ballungsgebiet gelegen ist (Lage) und ob sie durch Landwirtschaft, Industrie oder Handel, Dienstleistung und Verwaltung geprägt ist (Grad der Spezialisierung).

3. Empirische Prüfung, I. Versuch

3.1 Operationalisierung

Der erste Versuch, eine Wohnquartiersbeschreibung bei einer Repräsentativerhebung durch Interviewerbeobachtung zu erhalten, wurde von ZUMA 1979/80 bei der Erhebung zum ALLBUS 1980 unternommen. Dabei war von den Interviewern für jede angelaufene Adresse ein Kontaktprotokoll auszufüllen, unabhängig davon, ob ein Kontaktversuch zum Interview geführt hatte oder nicht. Diesem Kontaktprotokoll war eine Abfolge von sechs Fragen zur Beschreibung des Wohnquartiers der Zielperson angehängt. Durch diese Trennung der Wohnquartiersbeschreibung vom Erhebungsinstrument und dessen Ankoppelung an ein separat und für jede angelaufene Adresse auszufüllendes Kontaktprotokoll wurde sichergestellt, daß eine Wohnquartiersbeschreibung auch dann zur Verfügung steht, wenn der Interviewer nicht die Möglichkeit zur Befragung der Zielperson hatte

und deren soziologische Merkmale daher auch nicht durch Abfragen erhoben werden konnten.

Diese erste Wohnquartiersbeschreibung war kein völlig neues Instrument. Zum Erstellen der benötigten Fragenbatterie nahm man zunächst vorhandene, erprobte Fragen aus der Standarddemographie und ergänzte diese in Anlehnung an den oben aufgeführten Variablenkatalog. Im einzelnen gelangte dann folgende Fragenabfolge in den Kontaktprotokollanhang:

Eine Ortstypologie der Wohngemeinde mit Informationen zur Ortsgröße, zur Lage und zur Entwicklungsstufe des Ortes bzw. der Stadt:

1. Wie würden Sie den Ortstyp der Wohngemeinde bezeichnen?

Eingestuft werden soll nicht unbedingt die Gemeinde im Verwaltungssinn, sondern die mehr oder weniger geschlossene Siedlung, die Sie wahrnehmen

Einzelgehöft, Weiler	01
Dorf in rein ländlicher Umgebung	02
Dorf in der Nähe einer mittleren Stadt oder Großstadt	03
Ländliche Kleinstadt (bis 30.000 Einwohner)	04
Industrielle Kleinstadt (bis 30.000 Einwohner)	05
Stadt mittlerer Größe mit wenig Industrie (bis 100.000 Einwohner)	06
Stadt mittlerer Größe mit viel Industrie (bis 100.000 Einwohner)	07
Großstadt	08
Vorort einer Großstadt	09

Eine Feststellung der überwiegenden Landnutzung im Zielgebiet:

2. Wir haben nun einige Fragen zu der Wohngegend im unmittelbaren Umkreis des Zielhaushaltes. Bitte beantworten Sie diese Fragen nach dem Eindruck beim Besuch bzw. nach Ihrer eigenen Ortskenntnis.
Ist die Wohngegend ...

eher ein reines Wohngebiet	1
eher ein Mischgebiet mit ungefähr gleichem Anteil an Wohnungen und Geschäften bzw. Gewerbebetrieben	2
eher ein Geschäftszentrum (einschließlich Banken und Verwaltungsgebäuden) mit wenig Wohnungen	3
eher ein reines Gewerbe- bzw. Industriegebiet mit wenig Wohnungen	4

Eine Beschreibung der nächsten Umgebung des Zielhaushaltes, bestehend aus der Einordnung von Gebäudeart oder Siedlungsart:

3. Welche Art Häuser stehen unmittelbar im Umkreis des Zielhaushalts?
- | | |
|---|----|
| meist einzelne, freistehende Ein- oder Zweifamilienhäuser | 01 |
| meist aneinandergebaute Häuser/Reihenhäuser mit 1-2 Stockwerken | 02 |
| meist Häuser mit 3-4 Stockwerken, <u>teils</u> freistehend <u>teils</u> in Gruppen/Reihenhausbauweise | 03 |
| eine <u>durchgängige</u> Straßenrandbebauung mit 4-6 Stockwerken | 04 |
| kleinere Neubausiedlung (nach 1950 gebaut, kleinere oder größere Häuser, einschließlich Hochhäuser) | 05 |
| größere Neubausiedlung/Trabantenstadtsiedlung (nach 1950 gebaut) | 06 |
| keine Häuser | 07 |

Eine 10-er Skala zur Bewertung des Wohngebietes:

4. Wohngebiete kann man als besser oder schlechter einschätzen, wenn man die Wohnlage, die Qualität der Wohnungen und Häuser und die Leute beobachtet, die dort wohnen. Wie würden Sie auf dieser Skala die Wohngegend einschätzen, in der der Zielhaushalt liegt?

sehr gut												sehr schlecht
1	2	3	4	5	6	7	8	9	10			

Eine Klassifizierung des Haustyps, in dem der Zielhaushalt wohnt:

5. In welchem Haustyp wohnt der Zielhaushalt?
- | | |
|---|----|
| Bauernhaus | 01 |
| Freistehendes Einfamilienhaus (Bungalow, Siedlungshaus usw.) | 02 |
| Nicht-freistehendes Einfamilienhaus (Reihenhaus, Doppelhaus usw.) | 03 |
| Haus mit 2 bis 4 Parteien | 04 |
| Haus mit 5 bis 9 Parteien | 05 |
| Haus mit 10 bis 19 Parteien | 06 |
| Haus mit 20 und mehr Parteien | 07 |
| Notunterkunft, Baracke | 08 |
| Sonstiges, und zwar | 09 |

Eine Bestimmung der Nutzung des Gebäudes, in dem der Zielhaushalt wohnt:

6. Ist das Gebäude ein reines Wohnhaus oder sind darin auch Geschäfte, Büros, Warenlager etc. untergebracht?
- | | |
|---------------------|---|
| reines Wohnhaus | 1 |
| auch Geschäfte etc. | 2 |

3.2 Ergebnisse

Die aufgeführten sechs Fragen sind mit Hilfe eines theoretisch begründeten Variablenkatalogs formuliert worden und erscheinen leicht beantwortbar. Dennoch sind die hier gebrauchten Fragestellungen zu verwerfen, denn sie messen nicht das, was sie messen sollen: Ein Vergleich der Antworten mit den Ortsangaben der regionalen Erhebungseinheiten zeigt, daß eine richtige Zuordnung des zu Sehenden und des zu Erfahrenden wegen der Mehrdimensionalität der Antwortkategorien nicht möglich war. So wird etwa in Frage 1 sowohl nach der Ortsgröße als auch nach dem Industrialisierungsgrad gefragt. Bei der Einstufung der Kommunen wird die Auswertung dadurch erschwert, daß die Fragestellung von Frage 1 nicht eine Einstufung nach objektiven Kriterien verlangt, sondern nach der subjektiven Wahrnehmung einer mehr oder weniger geschlossenen Siedlungsform fragt. Hiermit ist - wie die nachfolgenden Beispiele zeigen - ein Interviewer überfordert. Ein Interviewer sieht Duisburg Stadt z.B. nicht als Teil der Agglomeration "Ruhr", sondern je nachdem wie gerade die unmittelbare Umgebung des Wohngebäudes des Zielhaushaltes aussehen mag, wird Duisburg mal als "Einzelgehöft", mal als "Dorf in rein ländlicher Umgebung", dann wieder als "Dorf in Stadtnähe" oder als "Stadt mittlerer Größe mit wenig Industrie", aber auch als "Vorort einer Großstadt" und manchmal auch als "Großstadt" angesehen.

Duisburg ist zwar das krasseste Beispiel, was auch auf die dort arbeitenden Interviewer zurückzuführen sein mag, aber es ist nicht das einzige Beispiel: So reicht auch die Einstufung von Versmold bei Bielefeld vom "Dorf in rein ländlicher Umgebung" bis "Großstadt", während Düsseldorf u.a. zur "ländlichen Kleinstadt" und die wendländische Kreisstadt Dannenberg zum "Dorf in rein ländlicher Umgebung" wird. Eine Überprüfung der von den Interviewern getätigten Ortstypzuordnungen für 25 "falsch" zugeordnete Kommunen/Orte durch das Befragungsinstitut GETAS, Bremen hat ergeben, daß die Interviewer nicht gefälscht haben, sondern nur das protokollierten, was sie subjektiv wahrgenommen oder vermutet haben. Wegen dieser sehr subjektiven Wahrnehmung gibt es ein Durcheinander bei den Zuordnungen "industriell" contra "ländlich" bzw. "mit wenig Industrie", jeweils abhängig davon, ob der Zielhaushalt mit Blick auf eine Fabrik oder mit Blick auf dünn besiedeltes und durchgrüntes Gebiet gesehen wurde. Ebenfalls durcheinander geht es mit den

Begriffen "Dorf" und "Stadt". Oft wurde dort ein "dörflicher Charakter" wahrgenommen, wo eigentlich "Mittelstadt" oder "Großstadt" stehen müßte. Oder es wurde - der Fragestellung folgend - ein nach der Eingemeindung einer Ortschaft in eine Großstadt noch erhaltengebliebener dörflicher Charakter als "Dorf" gesehen und bewertet. Auch mit einer Lagebestimmung wie "in rein ländlicher Umgebung" oder wie "in der Nähe einer mittleren Stadt oder Großstadt" können die Interviewer nichts anfangen. Denn was heißt Nähe? Bedeutet "Nähe" Sichtweite? Oder bedeutet "Nähe" eine Entfernung bis zu 5 km oder eine Entfernung bis zu 15 km?

Die überwiegende Nutzung in den Zielgebieten ist natürlich das reine Wohnen. Allerdings ist unplausibel, daß 78,8% aller erreichten Haushalte in reinen Wohngebieten gelegen sein sollen, und nur 19,0% aller Haushalte in Mischgebieten liegen. Aber selbst wenn im konkreten Fall diese Angaben die Realität widerspiegeln, so demonstrierten sie doch nur die Sinnlosigkeit dieser Frage. Denn es gelingt nicht, hiermit das Wohnumfeld der angelaufenen Haushalte hinsichtlich unterschiedlicher Nutzungsarten und "Gelegenheiten" zu beschreiben. Zudem zeigen sich Zuordnungsschwierigkeiten für die Interviewer, die z.B. jedes fünfte "Einzelgehöft" in einem "eher reinen Gewerbe- und Industriegebiet" gelegen sehen.

Die Beschreibung der näheren Umgebung des Standorts eines Zielhaushaltes hinsichtlich der Art der benachbarten Wohngebäude (oder der Siedlungsart) weist weitgehend eine plausible Zuordnung der auf Frage 3 gegebenen Antworten zu den bei Frage 1 angekreuzten Nennungen auf. Dieses darf aber nicht darüber hinwegtäuschen, daß auch hier zwei unterschiedliche Dimensionen gleichzeitig abgefragt wurden: Gebäudeart und Siedlungsart. Dieses sind Dimensionen, die sich nicht gegenseitig ausschließen und die einen Interviewer daher nur verunsichern. Auffällig ist, daß die meisten Städte nur aus maximal zwei unterschiedlichen Gebieten bestehen sollen - eine Beobachtung, die möglicherweise auf die geringe Größe der einzelnen "sample-points" des "ADM-Master-Samples" zurückzuführen ist. Dieses ist als Indiz zu werten, daß die Kategorien zur Gebäudeart nur ungenügend differenzieren.

In 3,6% aller Fälle läßt sich eine bei Frage 5 geleistete Zuordnung nicht plausibel der bei Frage 3 angekreuzten Antwort zuordnen. Allerdings wird der Gebäudetyp in Frage 5 sehr viel präziser abgefragt als es mit der

Gebäudeart in Frage 3 geschieht. Was nützt aber eine präzise Skala hinsichtlich Gebäudegröße, wenn die sich dahinter verbergenden Gebäudetypen über eine Größenangabe kaum oder gar nicht ausgemacht werden können?

Nur 7,3% der befragten Haushalte wohnen nach den Aussagen der Interviewer beim ALLBUS 1980 nicht in einem reinen Wohngebäude. Der ZUMABUS 1980 - ebenfalls eine repräsentative Querschnittsbefragung mit "ADM-Stichprobe" - erbrachte bei veränderter Fragenformulierung (siehe unten), zu diesem Komplex ein völlig anderes Ergebnis: 19,3% der befragten Haushalte wohnen nicht in einem reinen Wohngebäude.

Wie die sehr subjektive Zuordnung von individuellen Eindrücken zu den teils mehrdimensionalen, teils ungenügend differenzierenden und daher für eine objektive Quartiersbeschreibung unbrauchbaren Variablen vermuten läßt, ist der Rückschluß über eine Wohnquartiersbeschreibung auf soziologische Merkmale der in einem Wohnquartier dominanten Bevölkerungsgruppe mit dieser im ALLBUS 1980 eingesetzten Version einer Wohnquartiersbeschreibung noch nicht möglich. Um diesen Befund zu verdeutlichen, die Rückschlußwahrscheinlichkeit aufzuzeigen, wurde eine Reihe von Diskriminanzanalysen gerechnet. Sie zeigen, daß zwar noch eine relative Konsistenz zwischen den unterschiedlichen Antworten eines Interviewers hinsichtlich der Einschätzung von "Ortstyp", "Gebäudetyp Nachbarwohngebäude" und "Gebäudetyp Zielhaushalt" besteht. Inkonsistenzen zeigen sich jedoch da, wo vermutlich ähnliche Wohngebiete unterschiedlich gesehen werden. Daher ist eine schichtspezifische Besiedlung relativ homogener Bevölkerungsgruppen aus den vorliegenden Daten nicht herauszulesen.

Zu den Ergebnissen der Diskriminanzanalysen im einzelnen: In 31,8% der Fälle konnte über die Beschreibung des "Gebäudetyp Nachbarwohngebäude" der wahrgenommene "Ortstyp" richtig vorausgesagt werden (Tabelle 1, Diagramm 1).

In 33,8% aller Fälle war der "Gebäudetyp Nachbarwohngebäude" über den wahrgenommenen "Ortstyp" und den "Gebäudetyp Zielhaushalt" richtig vorauszusagen (Diagramm 2).

Nimmt man nun die "Stellung im Beruf" als Schichtvariable, so zeigt Diagramm 3, daß nur in 13% aller Fälle der "Gebäudetyp Nachbarwohnge-

bäude" (also der optische Eindruck des Quartiers) über die Schichtvariable richtig vorherzusagen war.

Bei nur 5,5% aller Fälle konnte über den "Ortstyp", den "Gebäudetyp Nachbarwohngebäude" und den "Gebäudetyp Zielhaushalt" auf die Schichtvariable richtig zurückgeschlossen werden (Diagramm 4). Dieser Meßwert wurde erzielt, nachdem die mit 28 Ausprägungen hoch differenzierte "Stellung im Beruf" (vgl. Pappi 1979: 84ff., 274ff., 280f.) über Gemeinsamkeiten der einzelnen Gruppen untereinander, auch hinsichtlich des von der Gruppe bevorzugten Wohnquartiers, auf elf Ausprägungen reduziert worden war. Bei 28 Ausprägungen dieser Variablen wurde die Schichtvariable über die genannten Quartiersvariablen nur in 2,7% aller Fälle richtig erschlossen.

Als Hauptfehlerquelle war allerdings zu berücksichtigen, daß sich die Interviewer mit der geforderten Art Kontaktprotokoll anscheinend etwas schwer taten. Denn es muß vermutet werden, daß manche Interviewer die Kontaktprotokolle erst längere Zeit nach dem Befragtenkontakt als Gedächtnisprotokoll ausgefüllt haben, ein Verfahren, daß sehr fehlerträchtig ist. Anders lassen sich eine Reihe von Unstimmigkeiten nicht erklären.

Die Mehrzahl der auftretenden Fehler muß freilich als eine falsche Einordnung des Gebietes gesehen werden, also eine nicht absichtliche Verfälschung des Wahrzunehmenden. Es wurde hier vom Gefühl und sicherlich nur mit einem halben Blick hinschauend, subjektiv richtig aber objektiv einwandfrei falsch zugeordnet.

Aufgrund dieser Ergebnisse wurde es notwendig, einen neuen Fragebogen zur Erfassung einer Wohnquartiersbeschreibung zu entwickeln, um eine objektive Beschreibung von Wohnquartieren zu ermöglichen. Die Erfahrungen mit diesem neuen Instrument werden in den folgenden Abschnitten dargestellt.

Tabelle 1: Vorhersagematrix für den "Ortstyp" über: "Gebäudetyp Nachbarwohngebäude" und "Gebäudetyp Zielhaushalt"

Code		Erscheinungsbild								
	1	Einzelgehöft, Weiler								
	2	Dorf in rein ländlicher Umgebung								
	3	Dorf in der Nähe einer mittleren Stadt oder Großstadt								
	4	Ländliche Kleinstadt								
	5	Industrielle Kleinstadt								
	6	Stadt mittlerer Größe mit wenig Industrie								
	7	Stadt mittlerer Größe mit viel Industrie								
	8	Großstadt								
	9	Vororte einer Großstadt								
		Vorausgesagte Gruppenzugehörigkeit								
		1	2	3	4	5	6	7	8	9
aktuelle Gruppenzugehörigkeit	1	81,0	14,3	4,8	0	0	0	0	0	0
	2	15,4	46,7	6,4	21,9	0	4,6	0	3,7	1,3
	3	10,3	32,5	11,5	29,3	0	12,1	0	4,3	0
	4	5,0	29,8	4,7	21,7	0	13,1	0	17,5	8,4
	5	2,6	21,1	2,6	21,6	0	12,3	0	29,1	10,6
	6	1,0	15,5	4,9	16,5	0	18,9	0	28,2	15,0
	7	2,2	12,4	4,0	17,8	0	7,1	0	46,2	10,2
	8	3,0	3,3	1,3	8,5	0	9,6	0	65,4	8,9
	9	7,6	11,4	1,4	24,2	0	11,8	0	34,9	8,7

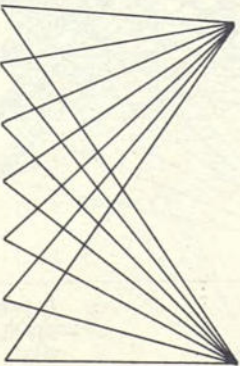
Insgesamt wurden 31,8% aller Fälle richtig zugeordnet.

Diagramm 1: Vorhersagematrix für den "Ortstyp" über: "Gebäudetyp Nachbarwohngebäude" und "Gebäudetyp Zielhaushalt"

vorherzusagende Variable		richtig zugeordnet (Treffergenauigkeit)	Indikatorvariablen	
Ortstyp	1 Einzelgehöft	81,0%	Gebäudetyp Nachbarwohngebäude	1 Einfamilienhaus, freistehend
	2 ländliches Dorf	46,7%		2 Reihenhaushaus, 1 - 2 stöckig
	3 stadtnahes Dorf	11,5%		3 Reihenhaushausbauweise, 3 - 4 stöckig
	4 ländliche Kleinstadt	21,7%		4 durchgängige Straßenrandbebauung
	5 industrielle Kleinstadt	0 %		5 kleinere Neubausiedlung
	6 wenig-industrielle Mittelstadt	18,9%		6 größere Neubausiedlung
	7 hoch-industrielle Mittelstadt	0 %		7 keine Nachbarhäuser
	8 Großstadt	65,4%	Gebäudetyp Zielhaushalt	1 Bauernhaus
	9 Vorstadt	8,7%		2 Einfamilienhaus, freistehend
				3 Einfamilienhaus, nicht freistehend
				4 Haus mit 2 - 4 Parteien
				5 Haus mit 5 - 9 Parteien
				6 Haus mit 10 - 19 Parteien
				7 Haus mit mehr als 20 Parteien
				8 Notunterkunft
				9 Sonstiges

Insgesamt wurden 31,8% aller Fälle richtig zugeordnet.

Diagramm 2: Vorhersagematrix für den "Gebäudetyp Nachbarwohngebäude" über: "Ortstyp" und "Gebäudetyp Zielhaushalt"

vorherzusagende Variable			Indikatorvariablen	
			Ortstyp	Gebäudetyp Zielhaushalt
Gebäudetyp Nachbarwohngebäude	1 Einfamilienhaus, freistehend	40,1%		1 Einzelgehöft
	2 Reihenhaushaus, 1 - 2 stöckig	36,6%		2 ländliches Dorf
	3 Reihenhaushausbauweise, 3 - 4 stöckig	18,7%		3 stadtnahes Dorf
	4 durchgängige Straßenrandbebauung	28,9%		4 ländliche Kleinstadt
	5 kleinere Neubausiedlung	25,2%		5 industrielle Kleinstadt
	6 größere Neubausiedlung	74,3%		6 wenig-industrielle Mittelstadt
	7 keine Nachbarhäuser	0 %		7 hoch-industrielle Mittelstadt
				8 Großstadt
				9 Vorstadt
				1 Bauernhaus
				2 Einfamilienhaus, freistehend
				3 Einfamilienhaus, nicht freistehend
				4 Haus mit 2 - 4 Parteien
				5 Haus mit 5 - 9 Parteien
				6 Haus mit 10 - 19 Parteien
				7 Haus mit mehr als 20 Parteien
				8 Notunterkunft
				9 Sonstiges

Insgesamt wurden 33,83% aller Fälle richtig zugeordnet.

Diagramm 3: Vorhersagematrix für den "Gebäudetyp Nachbarwohngebäude" über: "Stellung im Beruf" (11 Ausprägungen)

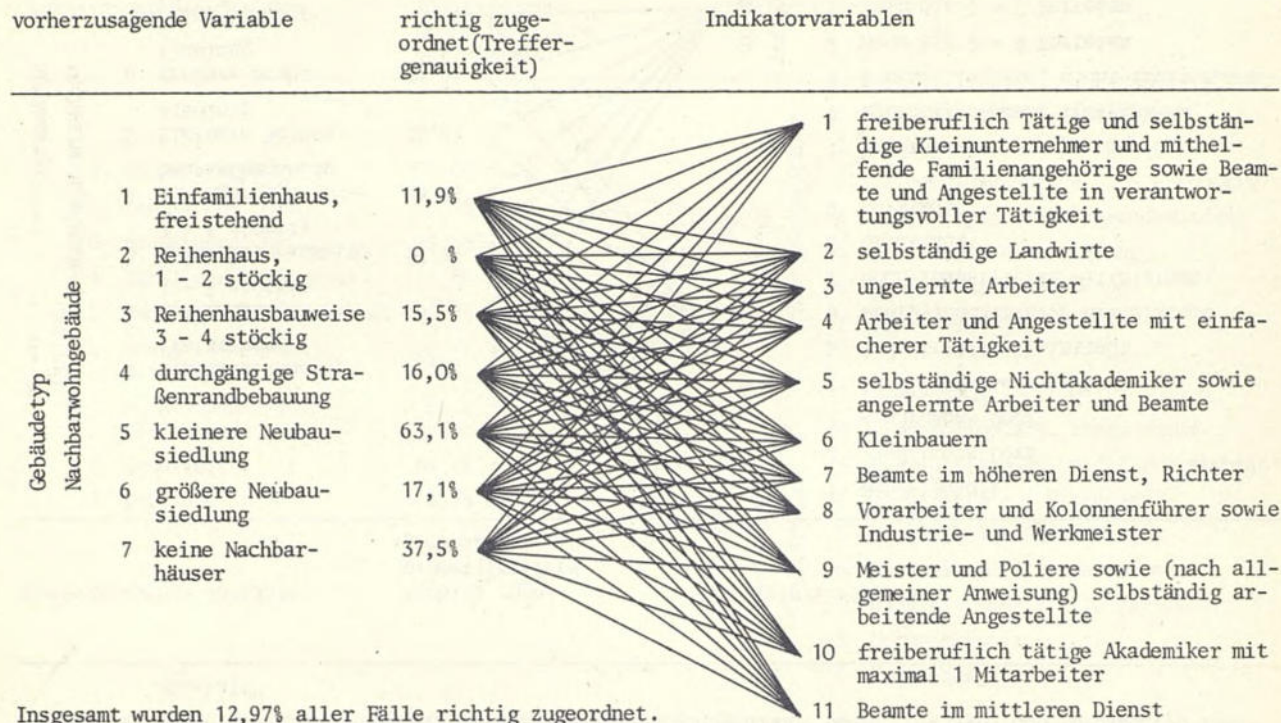
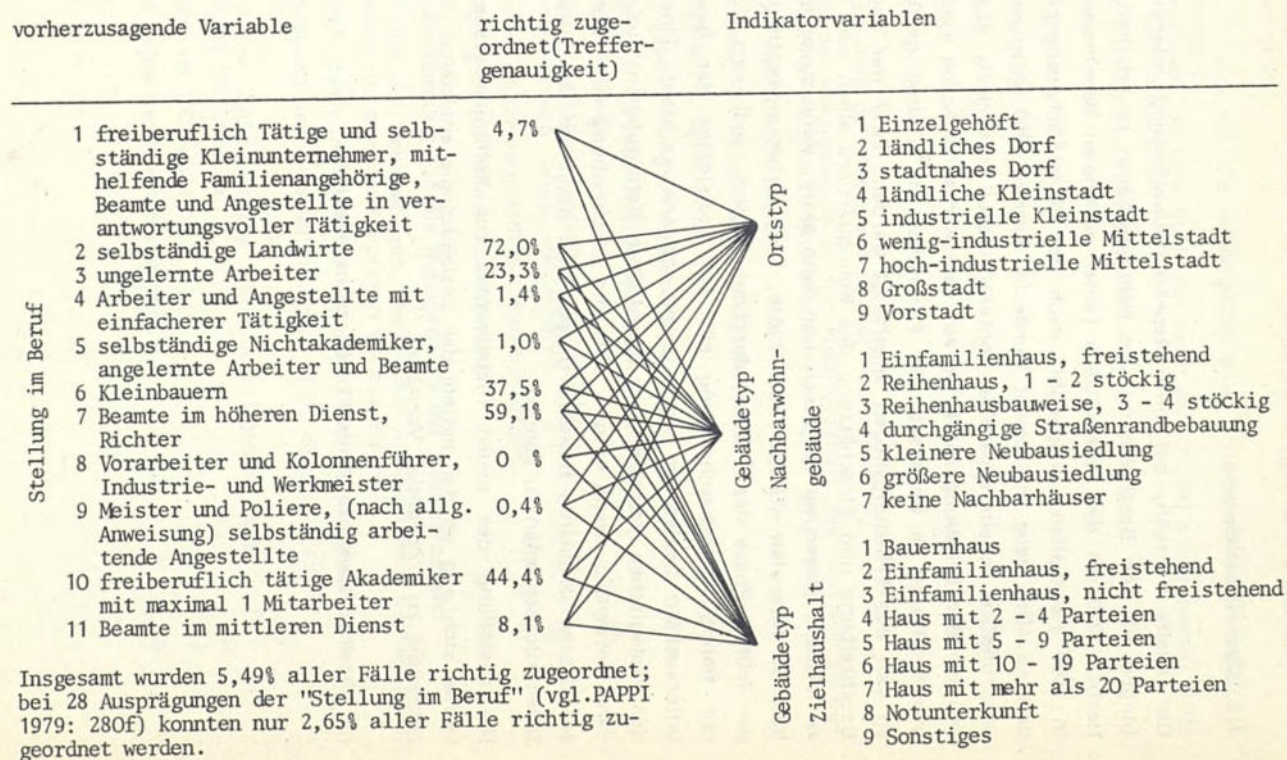


Diagramm 4: Vorhersagematrix für die "Stellung im Beruf" über: "Ortstyp", "Gebäudetyp Nachbarwohngebäude" und "Gebäudetyp Zielhaushalt"



4. Empirische Prüfung, II. Versuch

4.1 Operationalisierung

Der zweite Versuch, bei einer Repräsentativerhebung Interviewbeobachtungen für die Beschreibung von Wohnquartieren zu erhalten, wurde im Herbst 1980 mit dem ZUMABUS 5 (einer nationalen Mehrthemenbefragung in 2000 Haushalten, ausgewählt nach dem ADM-Stichprobenplan) unternommen. Das neue Instrument wurde im Sommer 1980 entwickelt und war ursprünglich in ein separates Kontaktprotokoll integriert. Aus Zeitgründen bei der Erhebung mußte es aus dem Kontaktprotokoll wieder herausgenommen und im Gegensatz zur Pretestversion in einer gekürzten Form an den Bus-Erhebungsbogen angehängt werden. Zwar war auch hier die Beschreibung und Einschätzung des Wohnquartiers allein vom Interviewer zu leisten. Dieser wurde jedoch nur dann aktiv, wenn zuvor im Zielhaushalt ein Interview stattgefunden hatte. Diese Erhebungsmethode schien in der frühen Phase des Testens durchaus möglich, weil noch ein Vergleich der Wohnquartiersbeschreibung mit den Sozialdaten der Bewohner der untersuchten Gebiete im Vordergrund der Analyse steht. Eine Trennung der ökologischen Variablen vom eigentlichen Befragungsinstrument ist erst dann notwendig, wenn es gilt, über die Beschreibung eines Wohnquartiers auf dessen dominante Bewohnergruppe und damit auf Merkmale von Ausfällen rückschließen zu können.

Die Erstellung des neuen Instrumentes zur Wohnquartiersbeschreibung lehnte sich eng an den mittels der Sozialökologie ermittelten Katalog der notwendig zu erhebenden Variablen.

Im einzelnen waren folgende Fragen zu beantworten:

Eine Bestimmung der Lage (als Entfernung) bezogen auf das (nächste) Großstadtzentrum:

1. Wie weit ist das nächste Großstadtzentrum von dem Haus der Zielperson entfernt?
(Eine Großstadt ist eine Stadt mit mehr als 100.000 Einwohnern). Bitte nicht in Luftlinie angeben!

Zielperson wohnt in Großstadtzentrum	01
bis 500 m	02
500 bis 1000 m	03
etwa 1 bis 2 km	04
etwa 2 bis 10 km	05
etwa 10 bis 25 km	06
weiter als 25 km	07

Eine Beschreibung der Wohngebäude der Nachbarschaft:

2. Wie sind die Wohngebäude in der Nachbarschaft des von der Zielperson bewohnten Hauses zu charakterisieren? (Nur eine Nennung!)

Ein- oder Zweifamilienhäuser	01
Reihenhäuser mit Vorgärten (1-2 Geschosse)	02
Reihenhäuser, 2-4geschossig	03
Reihenhausbewohnung, 4-6geschossig	04
Durchgängige Straßenrandbebauung, ohne Vorgärten, 4-6geschossig	05
Durchgängige Straßenrandbebauung mit vielen Hinterhäusern, engen Hinterhöfen	06
Hochhäuser (auch vermischt mit Reihenhäusern)	07
Keine Wohngebäude in der Nachbarschaft	08

Eine Altersbestimmung für die Wohngebäude der Nachbarschaft:

3. Wie alt sind die Wohngebäude, die in der unmittelbaren Nachbarschaft des Hauses, in dem die Zielperson wohnt, stehen?

Neubauten (nach dem Zweiten Weltkrieg), bis 35 Jahre alt	1
Etwa zwischen den Kriegen errichtet, d.h. 40 bis 60 Jahre alt	2
Ende letzten, Anfang dieses Jahrhunderts erbaut, d.h. 60 bis 120 Jahre alt	3
Viel mehr als 100 Jahre alt	4
Absolut keine Einschätzung möglich	5

Eine altersmäßige Zuordnung des Hauses des Zielhaushaltes zu den Gebäuden der Nachbarschaft:

4. Gehört das Haus, in dem die Zielperson wohnt, in die gleiche Altersgruppe?

ja	1
nein	2

Eine Beschreibung der Nutzung des Hauses des Zielhaushaltes:

5. Gibt es in dem Haus, in dem die Zielperson wohnt ... (Mehrfachnennungen möglich!)

- nur Wohnung(en)	1
- auch Laden und/oder Kneipe	1
- auch Büro und/oder Praxis (z.B. Arzt, Anwalt)	1
- auch Werkstatt (Handwerksbetrieb, kleine Fabrik)	1
- Stall, Scheune	1

Eine Beschreibung der Nutzungen in der Nachbarschaft:

6. Gibt es in unmittelbarer Nähe des Hauses, in dem die Zielperson wohnt... (Mehrfachnennungen möglich!)

- ein Einkaufszentrum	1
- eine Fabrik	1
- ein Büro-(hoch-)haus	1
- landwirtschaftlich genutzte Gebäude (Stall, Scheune, Schuppen für Maschinen u.ä.)	1

Ein Einschätzen der Nachbarhaushalte hinsichtlich deren Schichtzugehörigkeit:

7. Bitte schätzen Sie, zu welcher Schicht wohl die Mehrzahl der Nachbarn der Zielperson am ehesten zuzurechnen sind. Vergessen Sie hierbei aber nicht, daß die Zielperson sich in ihrer Schichtzugehörigkeit durchaus von der Mehrzahl der Nachbarn unterscheiden kann!

Arbeiterschicht	1
Mittelschicht	2
Obere Mittelschicht	3
Oberschicht	4

Fragen zur "Erreichbarkeit" und zur "Gebietsbewertung" fielen der oben erwähnten Kürzung zum Opfer.

Der "Ortstyp" brauchte im ZUMABUS 5 1980 nicht mehr abgefragt werden. Hier konnte die subjektive Interviewereinstufung durch ein objektiveres Instrument ersetzt werden: Der "Ortstyp" wird in Zukunft mit Hilfe eines

Stadt-Index (siehe Hoffmeyer-Zlotnik, 1981b) ermittelt werden. Die für die Konstruktion des Stadt-Index benötigten Informationen liefert das ADM-Ziehungsband für jede Primäreinheit (vgl. Schäfer 1979: 120ff.). Dieser Stadt-Index ordnet alle Städte auf einer 5er-Skala ein: a) nach dem Grad ihrer Spezialisierung, d.h., je nachdem, durch welchen der drei Wirtschaftssektoren eine Stadt geprägt wird, ob von der Landwirtschaft, der Industrie oder dem Dienstleistungssektor (vgl. Fourastie 1969); b) nach ihrer administrativen Größe und c) nach ihrer Lage in einem ländlichen Raum, in einem städtischen Ballungsgebiet bzw. in einer Agglomeration.

4.2 Ergebnisse

Auch mit dieser zweiten Version der Wohnquartiersbeschreibung ist noch kein voll zufriedenstellendes Prognoseinstrument für den Rückschluß von einer Wohnquartiersbeschreibung auf soziologische Merkmale dominanter Bewohnergruppen gelungen. Die Auswertung der Beobachtungsprotokolle aus dem ZUMABUS 5 1980 mit Hilfe einer Reihe von Diskriminanzanalysen zeigt deutlich, daß die hier getestete Version des Instrumentes zwar in einem höheren Maße Rückschlüsse auf soziologische Merkmale von Bevölkerungsgruppen zuläßt als es die erste Version aus dem ALLBUS 1980 möglich machte. Die Treffergenauigkeit läßt freilich immer noch zu wünschen übrig.

Zu den Diskriminanzanalysen im einzelnen:

Die Schichtvariable "Stellung im Beruf" (vgl. hierzu Diagramm 5) ist über Lage, Erscheinungsbild und Alter eines Wohnquartiers ("Entfernung Großstadtzentrum", "Gebäudetyp Nachbarwohngebäude", "Gebäudealter Nachbarwohngebäude") mit 13%iger Treffergenauigkeit vorausszusagen (Diagramm 6).

Die verwendete Schichtvariable ist der Standarddemographie entnommen (siehe Pappi 1979: 271) und bietet somit eine Vergleichsmöglichkeit zwischen der Selbsteinschätzung des Befragten - erhoben als Teil der Standarddemographie - und der Einschätzung der Nachbarn durch die Interviewer. Aus Gründen dieser Vergleichbarkeit soll die derzeitige Fassung

Diagramm 5: Vorhersagematrix für die "Stellung im Beruf" über: "Entfernung Großstadtzentrum", "Gebäudetyp Nachbarwohngebäude", "Gebäudealter Nachbarwohngebäude"

vorherzusagende Variable		richtig zugeordnet(Treffer- genauigkeit)	Indikatorvariablen	
Stellung im Beruf	1 freiberuflich Tätige und selbstständige Kleinunternehmer, mit-helfende Familienangehörige, Beamte und Angestellte in ver-antwortungsvoller Tätigkeit	16,9%		1 Zielperson wohnt in Großstadtzentrum
	2 selbstständige Landwirte	25,0%		2 bis 500 m
	3 ungelernte Arbeiter	0,0%		3 500 bis 1000 m
	4 Arbeiter und Angestellte mit einfacherer Tätigkeit	15,6%		4 etwa 1 bis 2 km
	5 selbstständige Nichtakademiker, angelernte Arbeiter und Beamte	4,3%		5 etwa 2 bis 10 km
	6 Kleinbauern	75,0%		6 etwa 10 bis 25 km
	7 Beamte im höheren Dienst, Richter	61,9%		7 weiter als 25 km
	8 Vorarbeiter und Kolonnenführer, Industrie- und Werkmeister	8,8%		1 Ein- oder Zweifamilienhäuser
	9 Meister und Poliere, (nach allg. Anweisung) selbstständig arbei-tende Angestellte	5,5%		2 Reihenhäuser mit Vorgärten (1 - 2 Geschosse)
	10 freiberuflich tätige Akademiker mit maximal 1 Mitarbeiter	16,7%		3 Reihenhäuser (2 - 4 Geschosse)
	11 Beamte im mittleren Dienst	2,9%		4 Reihenhausbauung (4 - 6 Geschosse)
		Gebäudetyp Nachbarwohn- gebäude	5 durchgängige Straßenrandbebauung, ohne Vorgärten (4 - 6 Geschosse)	
		Gebäudealter Nachbarwohn- gebäude	6 durchgängige Straßenrandbebauung mit vielen Hinterhäusern, enge Hinterhöfe	
			7 Hochhäuser (auch mit Reihenhäusern)	
			8 keine Wohngebäude in Nachbarschaft	
			1 Neubauten bis 35 Jahre alt	
			2 etwa zwischen den Kriegen errichtet	
Insgesamt wurden 12,66% der Fälle richtig zugeordnet (bei 11 Ausprägungen); bei 28 Ausprägungen von "Stellung im Beruf" konnten 3,82% der Fälle richtig zugeordnet werden.			3 Ende letzten, Anfang dieses Jahr-hunderts erbaut	
			4 viel mehr als 100 Jahre alt	
			5 absolut keine Einschätzung möglich	

Diagramm 6: Vorhersagematrix für "Gebäudetyp Nachbarwohngebäude" über: "Entfernung Großstadtzentrum" und "Gebäudealter Nachbarwohngebäude"

vorherzusagende Variable			Indikatorvariablen	
			richtig zugeordnet (Treffergenauigkeit)	
Gebäudetyp Nachbarwohngebäude	1 Ein- oder Zweifamilienhäuser	50,0%	Entfernung Großstadtzentrum	1 Zielperson wohnt in Großstadtzentrum
	2 Reihenhäuser mit Vorgärten (1 - 2 geschossig)	25,3%		2 bis 500 m
	3 Reihenhäuser, 2 - 4 geschossig	0 %		3 500 bis 1000 m
	4 Reihenhausbauung 4 - 6 geschossig	9,6%		4 etwa 1 bis 2 km
	5 durchgängige Straßenrandbauung, ohne Vorgärten, 4 - 6 geschossig	29,6%		5 etwa 2 bis 10 km
	6 durchgängige Straßenrandbauung mit vielen Hinterhäusern, enge Hinterhöfe	56,9%		6 etwa 10 bis 25 km
	7 Hochhäuser (auch vermisch mit Reihenhäusern)	18,6%		7 weiter als 25 km
	8 keine Wohngebäude in der Nachbarschaft	100,0%	Gebäudealter Nachbarwohngebäude	1 Neubauten (nach dem 2. Weltkrieg), bis 35 Jahre alt
				2 Etwa zwischen den Kriegen errichtet, d.h. 40 bis 60 Jahre alt
				3 Ende letzten, Anfang dieses Jahrhunderts erbaut, d.h. 60 bis 120 Jahre alt
				4 Viel mehr als 100 Jahre alt
				5 Absolut keine Einschätzung möglich

Insgesamt wurden 32,68% aller Fälle richtig zugeordnet.

der Schichtvariable der Standarddemographie nicht zur Diskussion gestellt werden.

Der "Gebäudetyp Nachbarwohngebäude" ist über "Entfernung Großstadtzentrum" und "Gebäudealter Nachbarwohngebäude" bei einem Drittel aller Fälle richtig zuzuordnen (Diagramm 7).

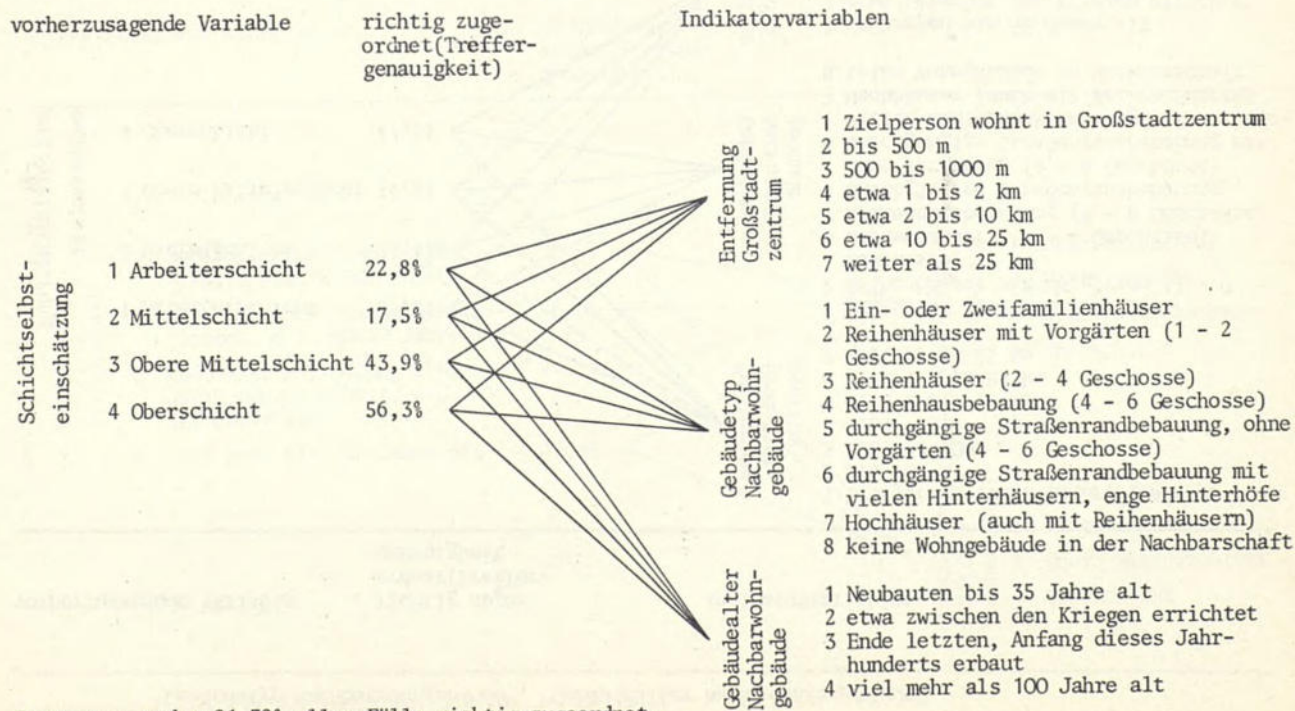
Die Schichtvariablen (Selbsteinschätzung, Fremdeinschätzung) sind, unter der Annahme der Richtigkeit oben angeführter Theorie, zu unpräzise, um einzelne Wohngebiete bestimmten sozialen Gruppen zuzuweisen. Ein Rückschluß auf die "Schichtselbsteinschätzung" über die Beschreibung des Gebietes (Diagramm 8) ist nur in 22% der Fälle möglich; ein Rückschluß auf die "Schichteinschätzung Nachbarschaft" durch den Interviewer über die Beschreibung des Gebietes (Diagramm 9) ist nur in 29% aller Fälle möglich.

"Gebäudealter Zielhaushalt" und Schichteinschätzung (sowohl Selbsteinschätzung als auch Fremdeinschätzung) hängen durchaus zusammen (Diagramm 10). Die hohe Anzahl der richtigen Zuordnungen ist jedoch trügerisch, weil sie stark abhängig ist von der hohen Anzahl der in Neubauten befragten Zielhaushalte. In 48% der Fälle konnte hier das "Gebäudealter Zielhaushalt" über die Schichteinschätzungen richtig zugeordnet werden. Versucht man allerdings einen Rückschluß auf das Gebäudealter, in dem die Befragungsperson wohnt, über befragungspersonbezogene Schichtvariablen wie "Schulabschluß des Befragten", "Ausbildungsabschluß des Befragten" und "Stellung im Beruf", so ist nur noch in 27% der Fälle eine richtige Zuordnung möglich.

Zusammenfassend ist festzustellen:

Schwächen des neuen Erhebungsinstruments bestehen in zu großen Spielräumen, die dem Interviewer für Ermessungsentscheidungen bei der Zuordnung zu den Antwortvorgaben eingeräumt wurden. Die Idee dieser "weichen" Antwortvorgaben ist es gewesen, dem Interviewer die Angst vor seinen erwartungsgemäß geringen baugeschichtlichen Kenntnissen zu nehmen und ihn keinem unmäßigen Leistungsdruck auszusetzen. Daß die Antwortvorgaben trotz der weiten Spielräume hinreichend differenzieren würden, konnte nur erhofft werden. Wie die Erfahrung später zeigte, scheint dieses, vor allem bezüglich der Bestimmung der Gebäudetypen, nicht der Fall zu sein. So muß für die Zukunft gerade die Bestimmung

Diagramm 7: Vorhersagematrix für "Schichtselbsteinschätzung" über: "Entfernung Großstadtzentrum", "Gebäudetyp Nachbarwohngebäude" und "Gebäudealter Nachbarwohngebäude"



Insgesamt wurden 21,72% aller Fälle richtig zugeordnet.

Diagramm 8: Vorhersagematrix für "Schichteinschätzung Nachbarschaft" über: "Entfernung Großstadtzentrum",
 "Gebäudetyp Nachbarwohngebäude", "Gebäudealter Nachbarwohngebäude"

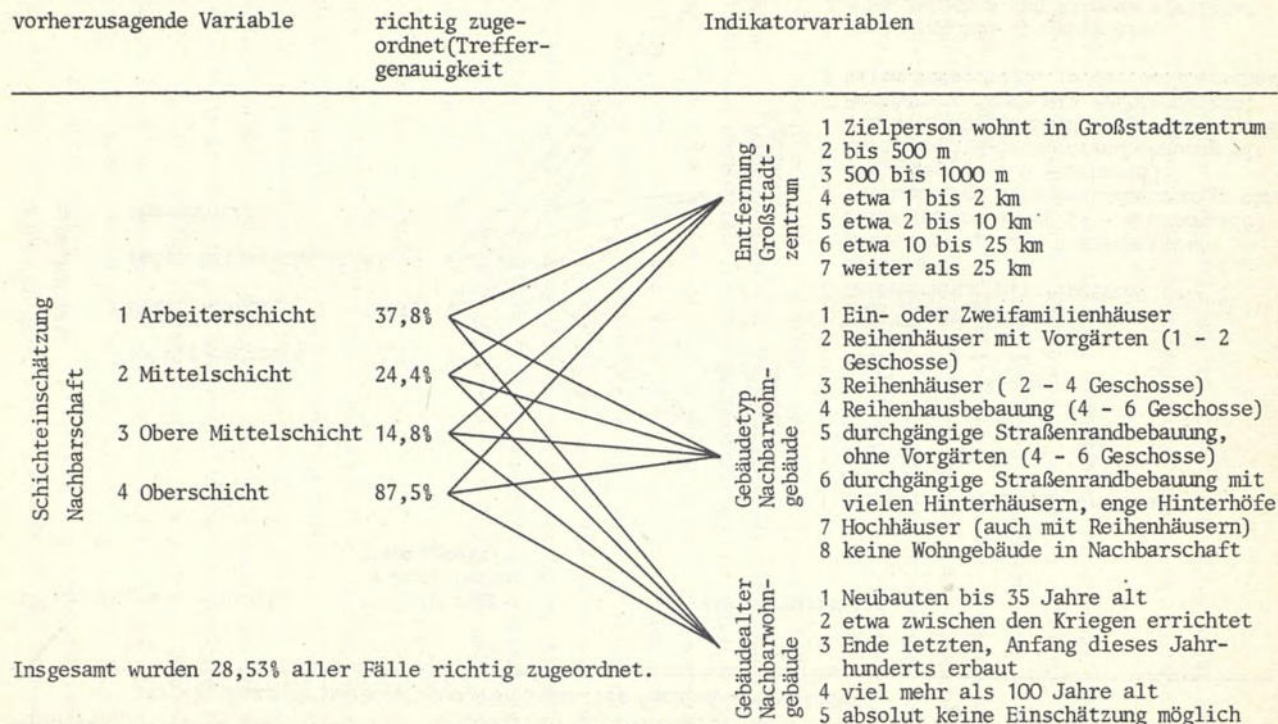


Diagramm 9: Vorhersagematrix für "Gebäudealter Zielhaushalt" über: "Schichtselbsteinschätzung" und "Schichteinschätzung Nachbarschaft"

vorherzusagende Variable		richtig zugeordnet (Treffergenauigkeit)	Indikatorvariablen	
Gebäudealter Zielhaushalt	1 Neubauten (nach dem 2. Weltkrieg), bis 35 Jahre alt	60,4%	Schichtselbst- einschätzung	1 Arbeiterschicht
	2 etwa zwischen den Kriegen errichtet, d.h. 40-60 Jahre alt	0,0%		2 Mittelschicht
	3 Ende letzten, Anfang dieses Jahrhunderts erbaut, d.h. 60-120 Jahre alt	42,1%		3 Obere Mittelschicht
	4 viel mehr als 100 Jahre alt	13,6%		4 Oberschicht
			Schichtein- schätzung Nachbarschaft	1 Arbeiterschicht
				2 Mittelschicht
				3 Obere Mittelschicht
				4 Oberschicht

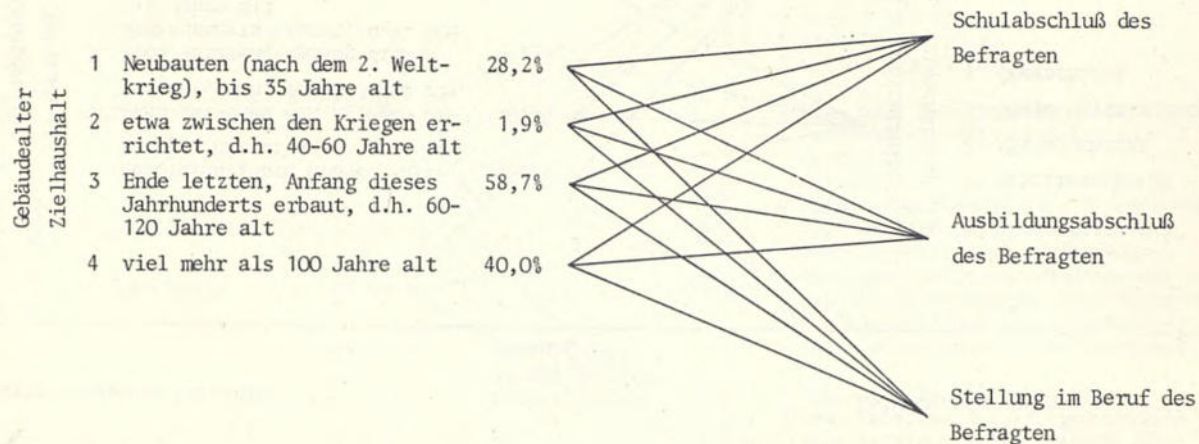
Insgesamt wurden 47,95% aller Fälle richtig zugeordnet.

Diagramm 10: Vorhersagematrix für "Gebäudealter Zielhaushalt" über: "Schulabschluß des Befragten", "Ausbildungsabschluß des Befragten", "Stellung im Beruf des Befragten"

vorherzusagende Variable

richtig zugeordnet (Treffergenauigkeit)

Indikatorvariablen



Insgesamt wurden 27,14% der Fälle richtig zugeordnet.

der Gebäudetypen kategorienmäßig stärker präzisiert und härter auf das Sichtbare reduziert werden. Außerdem wird, wie sich wider Erwarten herausgestellt hat, eine Gebäudetypbeschreibung für das Wohngebäude des Zielhaushaltes notwendig, denn schon bei 13% aller Fälle gehört das Wohngebäude des Zielhaushaltes nicht mehr der gleichen Altersgruppe an, der die die Nachbarschaft prägenden Wohngebäude zuzurechnen sind.

5. Schlußfolgerungen

Die Ergebnisse beider Versuche, über eine Wohnquartiersbeschreibung auf soziologische Merkmale von Quartiersbewohnern rückzuschließen, machen deutlich, daß das vorliegende Instrument zwar den richtigen Weg weist, das Ziel aber noch nicht erreicht ist. Als nächsten Schritt zu einer ausgereiften Wohnquartiersbeschreibung werden wir eine kleine Gruppe von Interviewern in einem nachkontrollierbaren und räumlich überschaubaren Feld einsetzen. Hierbei sollen dann sowohl Beobachtungsexperimente als auch Einsätze unter normalen Feldbedingungen durchgeführt werden.

In einer solchen Versuchsreihe muß und wird sich dann erweisen, welcher Art eine Präzisierung der Fragen und deren Antwortvorgaben zu sein hat und welches Ausmaß an Verfälschungen durch die Subjektivität der Betrachtung und des Beschreibens durch Interviewer (und nicht durch extra geschulte Beobachter) hingenommen werden muß. Auch wird sich zeigen, welche Verzerrungen als Folgen möglicher Nachlässigkeit beim Ausfüllen des Beobachtungsbogens als unvermeidbare Fehlerquoten hinzunehmen sein werden. Eine solche Testreihe muß dann zeigen, mit welcher Trefferwahrscheinlichkeit ein richtiger Rückschluß über eine Wohnquartiersbeschreibung auf die in diesem Quartier dominante Bewohnergruppe gelingen wird. Erst im Anschluß an solch eine Testreihe ist eine Entscheidung möglich, ob auf die soziologischen Merkmale der in einem städtischen Teilgebiet dominierenden Bewohnergruppe über eine Beschreibung von deren Wohngebieten rückgeschlossen werden kann oder ob die Subjektivität der Beschreibenden eine hohe Rückschlußwahrscheinlichkeit unmöglich macht.

Literatur

- Burgess, Ernest W., 1925: The Growth of the City: An Introduction to a Research Project. S.47-62 in: Robert E. Park, Ernest W. Burgess und Roderick D. McKenzie (Hrsg.), The City. Chicago, London: University of Chicago Press.
- Burgess, Ernest W., 1928: Urban Areas. S.113-138 in: T.V. Smith und Leonard White (Hrsg.), Chicago: An Experiment in Social Science Research. Chicago: Greenwood Press.
- Duncan, Otis D. und Beverly Duncan, 1957: The Negro Population of Chicago: A Study of Residential Succession. Chicago, London: University of Chicago Press.
- Fourastie, Jean, 1969: Die grosse Hoffnung des 20. Jahrhunderts. Köln: Bund.
- Friedrichs, Jürgen, 1977: Stadtanalyse: Soziale und räumliche Organisation der Gesellschaft. Reinbek: Rowohlt.
- Friedrichs, Jürgen (Hrsg.), 1978: Stadtentwicklungen in kapitalistischen und sozialistischen Ländern. Reinbek: Rowohlt.
- Hamm, Bernd, 1977: Die Organisation der städtischen Umwelt. Frauenfeld, Stuttgart: Huber.
- Hatt, Paul, 1946: The Concept of Natural Area. American Sociological Review, 11: 423-427.
- Hoffmeyer-Zlotnik, Jürgen H.P., 1977: Gastarbeiter im Sanierungsgebiet. Hamburg: Christians.
- Hoffmeyer-Zlotnik, Jürgen H.P., 1979: Eine Analyse des sozialökologischen Prozesses der Bevölkerungssukzession. S.114-136 in: Bernd Hamm (Hrsg.), Lebensraum Stadt. Frankfurt, New York: Campus.
- Hoffmeyer-Zlotnik, Jürgen H.P., 1981a: Wohnquartiersbeschreibung als Mittel zur Messung soziologischer Merkmale von Ausfällen. ZUMA-Nachrichten, 8: 5-24.
- Hoffmeyer-Zlotnik, Jürgen H.P., 1981b: Zur Konstruktion eines neuen Stadt-Index. ZUMA-Nachrichten, 9: 47-52.
- Pappi, Franz Urban, 1979: Sozialstrukturanalyse mit Umfragedaten. Königstein/Ts.: Athenäum.
- Schäfer, Felix, 1979: Muster-Stichproben-Pläne. München: Moderne Industrie.
- Theodorsen, George A. (Hrsg.), 1961: Studies in Human Ecology. Evanston, New York: Harper & Row.
- Zapf, Katrin, 1969: Rückständige Viertel. Frankfurt/M.: Europäische Verlagsanstalten.

EINE SOZIO-DEMOGRAPHISCHE BESCHREIBUNG DER TRÄGER POST-MATERIALISTISCHER EINSTELLUNGEN

Manfred Küchler

1. Einleitung

Zielsetzung dieses Beitrags ist es, anhand einer im Bereich der politischen Soziologie gut eingeführten Fragestellung Methode und Technik eines multivariaten Analyseverfahrens zu demonstrieren. Diese methodologische Schwerpunktsetzung rechtfertigt es, hier inhaltliche Ergebnisse zu präsentieren, die bis zum Zeitpunkt des Erscheinens dieses Sammelbandes schon allgemein akzeptiertes Wissen einschlägig forschender Sozialwissenschaftler geworden sein dürften. Es geht also weniger um die spezifischen inhaltlichen Ergebnisse als vielmehr das Prinzipielle an der Art und Weise, wie derartige Fragestellungen behandelt werden können. Der Neuigkeitswert dieses Beitrags liegt insbesondere in der anwendungsbezogenen Demonstration eines vor erst wenigen Jahren entwickelten statistischen Verfahrens der schrittweisen Auswahl von Erklärungsfaktoren bei nicht-metrischer Datenkonstellation, das ein funktionales Äquivalent zur bekannten und in gängigen Programmpaketen (etwa SPSS) implementierten schrittweisen multiplen Regression bei metrischen Merkmalen darstellt.

2. Inhaltliche Problemstellung

Ronald Inglehart (1977) hat in den siebziger Jahren eine "stille Revolution" entdeckt, eine fundamentale Verschiebung von "materiellen" hin zu "post-materiellen" Werten. Unter Rückgriff auf die Maslowsche Bedürfnishierarchie wurde die Herausbildung eines post-materiellen Wertesystems zunächst mit zunehmendem Wohlstand erklärt, dann angesichts ökonomischer Krisen und Fortdauer des Syndroms als dauerhaftes neues sozio-kulturelles Muster gesehen. Wir wollen an dieser Stelle darauf verzichten, die theoretische Diskussion um diesen Begriff zu referieren. Festzustellen bleibt, daß die von Inglehart vorgeschlagene Operationalisierung Eingang in eine Vielzahl von empirischen Studien gefunden hat, so vor allem die

große international vergleichende "Political Action"-Studie (Barnes und Kaase 1979) und eben auch die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS 1980).

Wir vertreten die Auffassung, daß bislang eine empirische Klärung des mit dem Etikett "Postmaterialismus" versehenen Syndroms aussteht, oder anders formuliert: daß die Gültigkeit der Operationalisierung dieses Begriffs noch nicht etabliert werden konnte¹. Dies gilt unabhängig davon, ob, wie im ALLBUS, die 4-Item-Kurzform oder die Langfassung der "Political Actions"-Studie (Barnes und Kaase 1979: 564ff.) benutzt wird. Unser Gültigkeitsanspruch reduziert sich dabei nicht auf eine bloße Übereinstimmung von theoretischem Konstrukt und operationaler Meßvorschrift, die u.U. durch Rekurs auf den Typus "operationale Definition" per fiat hergestellt werden könnte ("Postmaterialismus ist das, was die vorgegebenen Items messen"), sondern ist abgeleitet aus der Forderung, daß Sozialwissenschaftler das Denken und Handeln der Individuen in den Kategorien beschreiben müssen, deren sich die Individuen selbst bedienen, die also den Praktiken der Gesellschaftsmitglieder homomorph sind. Vorausgesetzt sie haben die Absicht, soziale Prozesse so zu beschreiben, daß daraus strukturelle Prognosen abgeleitet werden können². Diese Forderung kann jedoch allein mit Hilfe konventioneller Forschungsstrategien (Umfragen mit standardisierten Instrumenten) nicht eingelöst werden. Vielmehr müssen solche repräsentativ angelegten Studien, deren Hauptanliegen die Ermittlung der Verteilung bestimmter Einstellungen ist, ergänzt werden durch Forschungen, die aufgrund ihres Designs eher in der Lage sind, die spezifische Ausformung bestimmter Einstellungen und Prozesse der Bildung und Veränderungen derartiger Haltungen zu erfassen³. Derartige Designs müssen sich aufgrund forschungsökonomischer Beschränkungen zumeist mit sehr geringen Fallzahlen unter Preisgabe des Anspruchs auf statistische Repräsentativität begnügen. Mithin kommt repräsentativen Umfragen die wichtige Aufgabe eines "screening device" zu, damit darauf aufbauende "qualitative Studien" strategisch möglichst günstig platziert werden können. Wir setzen uns hier damit inhaltlich das Ziel, zu einer möglichst präzisen Beschreibung von "Postmaterialismus" aufgrund äußerlich bestimmbarer Merkmale zu gelangen, also solcher Merkmale, die auch ohne größeren Zusatzaufwand als Kriterium für die Abgrenzung einer Untersuchungsgesamtheit dienen können, d.h. im wesentlichen sozio-demographische Merkmale.

3. Überblick über die Analyseschritte

Während es klassische metrische Verfahren (Regressionsanalyse, Faktorenanalyse) erlauben, relativ große Mengen an Erklärungsfaktoren (Merkmale) simultan zu verarbeiten und mit Hilfe statistischer Kriterien die wesentlichen aussondern, erfordern die neueren nichtmetrischen Verfahren der multivariaten Analyse in der Regel eine Vorauswahl der zu betrachtenden unabhängigen Variablen. Ausgangspunkt dieser Verfahren ist nämlich eine mehrdimensionale Häufigkeitsverteilung⁴, deren Zellenbesetzungen gewisse Mindestgrenzen nicht unterschreiten dürfen. Somit besteht eine Restriktion hinsichtlich der Zahl der simultan zu betrachtenden unabhängigen Merkmale und auch hinsichtlich der jeweils unterschiedenen Ausprägungen; jeweils in Abhängigkeit von der Zahl der zur Verfügung stehenden Fälle (Untersuchungseinheiten).

Gewöhnlich wird die aus inhaltlichen Gründen vorgenommene Einschränkung (hier auf sozio-demographische Faktoren) auch bei stärkerer Gruppierung der Ausprägungen noch nicht zu einer Zellenzahl (in der mehrdimensionalen Häufigkeitstabelle) führen, die in Anbetracht der Gesamtzahl der Fälle eine angemessene Besetzung gewährleistet. An dieser Stelle ist es nach einer inhaltlichen Vorauswahl sinnvoll, ein durch statistische Kriterien gesteuertes Auswahlverfahren einzusetzen, das aus der Menge aller inhaltlich interessanten Merkmale diejenige Menge von unabhängigen Variablen herausfiltert, die insgesamt die Verteilung der betrachteten abhängigen oder Zielvariablen am stärksten beeinflusst. Wir werden diesen Prozeß unten im einzelnen darstellen.

Nach der Bestimmung der einflußstärksten Menge von unabhängigen Merkmalen kann dann eine detaillierte Analyse dieser Datenkonstellation erfolgen. Dieser zweite Schritt hat die Funktion

- augenfällige Regelmäßigkeiten der mehrdimensionalen Häufigkeitsverteilung statistisch abzusichern, und so
- ein möglichst einfaches, aber dennoch den empirischen Daten hinreichend angepaßtes Modell zu entwickeln bzw. ein solches a priori Modell zu überprüfen, und damit
- eine relative Ausdifferenzierung des Gesamteinflusses zu leisten.

Für diesen zweiten Schritt sind in den letzten 10 oder 15 Jahren mehrere Ansätze entwickelt worden, die allesamt nur geringe und damit realitätsgerechte Ansprüche an das Meßniveau der Daten stellen (vgl. hierzu Küchler 1979 bzw. Küchler und Schwedler 1980). Wir werden hier den auf Grizzle, Starmer und Koch zurückgehenden nach den Anfangsbuchstaben der Autorennamen kurz GSK-etikettierten Ansatz verwenden⁵.

4. Daten

Wie schon erwähnt, ist im ALLBUS 1980 die Kurzform der Postmaterialismus-Operationalisierung gewählt worden. Den Befragten werden dabei vier Items vorgelegt, unter denen sie ihre ersten drei Präferenzen⁶ hinsichtlich der Vorgabe, welche der folgenden Ziele in der Politik der/die Befragte persönlich für am wichtigsten hält, nennen sollen. Diese Ziele sind wie folgt vorgegeben:

- A - Aufrechterhaltung von Ruhe und Ordnung in diesem Land
- B - Mehr Einfluß der Bürger auf die Entscheidungen der Regierung
- C - Kampf gegen die steigenden Preise
- D - Schutz des Rechts auf freie Meinungsäußerung

Als "Postmaterialisten" werden bei dieser Operationalisierung diejenigen definiert, deren erste zwei Präferenzen BD oder DB lauten, während "Materialisten" sich durch Wahl von AC oder CA zu erkennen geben. Die übrigen Befragten gehören zu einem Mischtypus. Geleitet von der inhaltlichen Überlegung, daß instabile wirtschaftliche Verhältnisse auch durchaus "postmaterialistisch" gesonnenen Befragten eine schlechte Vorbedingung für die Verwirklichung einer "neuen Politik" (z.B. in Fragen des Umweltschutzes) sind, also eine Präferenz für "C" qualitativ zu unterscheiden ist von dem reinen "law and order"-Item A, werden zunächst zwei unterschiedliche Mischtypen betrachtet. Wird Item C zusammen mit B oder D genannt, so sollte dies einen eher zum Postmaterialismus neigenden Mischtypus konstituieren im Vergleich zu einer Präferenz für Item A. Diese Abgrenzungen führten zu folgender Verteilung⁷:

Tabelle 1: Häufigkeit der einzelnen Typen

1) Postmaterialist	(BD, DB)	390
2) PM neigender Mischtyp	(BC, CB, DC, CD)	491
3) M neigender Mischtyp	(AB, BA, AD, DA)	921
4) Materialist	(AC, CA)	1101
Summe		2903
	Fehlende Werte	52

Die möglichen Erklärungsfaktoren sind in Tabelle 2 dargestellt⁸:

Tabelle 2: Erklärungsfaktoren

Merkmal	Ausprägungen	Anzahl
Alter	a) 18-25, 25-35, 35-45, 45-60, 60- b) 18-35, 35-60, 60 und darüber	5 3
subjektive Schicht-einstufung	Unter-Arbeiterschicht/Mittelschicht/ ob. Mittelschicht-Oberschicht	3
Schulbildung	- Hauptschulabschluß/mittlere Reife- Fachhochschule/Abitur, Hochschule	3
Religionszugehörig-keit	evangelisch/katholisch/sonstiges	3
Kirchgangshäufig-keit	bis einmal im Monat/mehrmals im Jahr/ seltener oder nie	3
Geschlecht	männlich/weiblich	2

Erste, zweidimensionale Analysen ergaben weiterhin, daß die inhaltlich abgeleitete Differenzierung der Mischtypen sich in dem empirischen Daten nicht wiederfindet, so daß im weiteren diese beiden Mischtypen zu einer Kategorie zusammengefaßt werden⁹. Die so entstandene Variable mit den Ausprägungen 'Postmaterialist', 'Mischtyp' und 'Materialist' werden wir im folgenden kurz mit PM bezeichnen.

Auswahlverfahren

Nach dem von Higgins und Koch (1977) vorgeschlagenen Verfahren werden in einem ersten Schritt zunächst alle unabhängigen Variablen der nach inhaltlichen Gesichtspunkten abgegrenzten Liste (die "Kandidaten"-Variablen) mit der Zielvariablen kreuztabelliert und das übliche Chi-Quadrat-Maß nach Pearson berechnet. Da die einzelnen unabhängigen Variablen in der Regel verschieden viele Ausprägungen haben - hier zwischen 2 und 5 - wird dieser Chi-Quadrat-Wert durch die Anzahl der Freiheitsgrade dividiert. Diese Zahl dient als Kriterium zur Bestimmung der einflußstärksten unabhängigen Variablen. Wir dokumentieren die Ergebnisse des ersten Schritts in der nachfolgenden Tabelle:

Tabelle 3: Schritt 1 des Auswahlverfahrens

PM mit	Chi ²	df	Chi ² /df
ALTER (3)	249.1	4	62.3
ALTER (5)	281.8	8	35.2
SCHULBILDUNG (3)	336.3	4	84.1
RELIGION (3)	30.7	4	7.7
KIRCHGANG (3)	52.4	4	13.1
GESCHLECHT (2)	22.2	2	11.1
SCHICHT (3)	108.8	4	27.2

Die Schulbildung erweist sich als die mit Abstand bedeutsamste unabhängige Variable. Da der Chi-Quadrat-Wert signifikant ist, wird diese Variable ausgewählt.

Im zweiten Schritt geht es uns darum, diejenigen der Kandidaten-Variablen (Ausgangsliste einer schon ausgewählten Variablen) zu ermitteln, die zusammen mit der schon ausgewählten Variablen die höchste Assoziation mit der Zielvariablen zeigt. Dazu werden Kombinationsvariablen gebildet, die jeweils eine der Kandidatenvariablen kreuzt mit der schon ausgewählten. Beispielsweise führt dies bei der trichotomen Altersvariablen zu einer Kombinationsvariablen mit $9 = 3 \times 3$ Ausprägungen, die alle Kom-

binationen von Altersgruppen und Schulbildung wiedergeben. Mit diesen Kombinationsvariablen werden nun wieder - formal gesehen zweidimensionale - Kreuztabellierungen mit der Zielvariable erstellt und wie zuvor die Größe Chi-Quadrat dividiert durch Freiheitsgrade berechnet. Die Ergebnisse im einzelnen sind wieder in einer Tabelle zusammengestellt:

Tabelle 4: Schritt 2 des Auswahlverfahrens

PM mit Kombi-Variablen aus SCHULBILDUNG und	Chi ² / df	Summe part. Chi ²
ALTER (3)	539.4 / 16 = 33.7	-
ALTER (5)	570.1 / 28 = 20.4	-
RELIGION (3)	358.7 / 16 = 22.4	-
KIRCHE (3)	352.1 / 16 = 22.0	-
GESCHLECHT (2)	364.9 / 10 = <u>36.5</u>	27.4+0.2+3.2=30.8 (6)
SCHICHT (3)	343.6 / 16 = 21.5	-

Es erweist sich, daß in Kombination mit der als einflußstärksten Variablen ermittelten SCHULBILDUNG das GESCHLECHT den Gesamteinfluß maximiert¹⁰. Damit ist aber noch nicht gesagt, daß dieser relativ gesehen beste Zuwachs an Gesamterklärung nun auch absolut gesehen bedeutsam, d.h. hier statistisch signifikant ist. Ab diesem Schritt tritt ein weiteres Abbruchkriterium in Aktion, indem die partielle Assoziation der neuen Kandidatenvariablen mit der Zielvariablen unter Kontrolle der schon ausgewählten Variablen bestimmt wird. Dies kann in einfacher Weise dadurch geschehen, daß das Pearsonsche Chi-Quadrat über alle Partialtafeln aufsummiert und in Relation zu der Summe der Freiheitsgrade auf Signifikanz überprüft wird¹¹. In unserem Fall führt diese Überprüfung zu einem signifikanten Resultat. Anderenfalls wäre die zweitbeste Kandidatenvariable nach dem Chi-Quadrat/df-Kriterium - hier also die trichotome Altersvariable - überprüft worden, usw. Erreicht keine der verbliebenen Kandidatenvariablen Signifikanz, so bricht der Auswahlprozeß ab.

Der dritte Schritt erfolgt ganz analog zum zweiten, nur die Kombinationsvariablen werden komplexer, da nun schon zwei Variablen ausgewählt sind.

Die Ergebnisse auf einen Blick:

Tabelle 5: Schritt 3 des Auswahlverfahrens

PM mit Kombi-Variablen aus SCHULBILDUNG, GE- SCHLECHT und	Chi ²	(df)
ALTER (3)	<u>571.9</u>	(34)
RELIGION (3)	394.4	(34)
KIRCHGANG (3)	378.7	(34)
SCHICHT (3)	382.0	(34)

Der Einfachheit halber haben wir die fünfstufige Altersvariable bereits fortgelassen. Damit haben alle Kombinationsvariablen $3 \times 2 \times 3 = 18$ Ausprägungen, die 18×3 -Tabelle also $17 \times 2 = 34$ Freiheitsgrade, somit ändert eine Division mit der Zahl der Freiheitsgrade nicht die Rangfolge. Die trichotome Altersvariable erweist sich nun relativ am einflußstärksten. Für die nun insgesamt sechs Partialtafeln mit je $2 \times 2 = 4$ Freiheitsgraden ergibt sich ein aufsummiertes Chi-Quadrat von 187.5, das bei $df = 24$ auf dem 0.1%-niveau signifikant ist.

Nun gleich die Ergebnisse des vierten Schritts in einer Tabelle:

Tabelle 6: Schritt 4 des Auswahlverfahrens

PM mit Kombi-Variablen aus SCHULBILDUNG, GE- SCHLECHT, ALTER und	Chi ² /(df)
RELIGION (3)	640.6 / 106 = 6.04
KIRCHGANG (3)	679.2 / 106 = 6.41
SCHICHT (3)	617.7 / 102 = 6.06

Die Kombinationsvariablen haben theoretisch $3 \times 2 \times 3 \times 3 = 54$ Ausprägungen, die Kreuztabelle mit PM somit $53 \times 2 = 106$ Freiheitsgrade. Die niedrige Zahl von 102 bei Schicht zeigt an, daß nicht alle Ausprägungen der Kombinationsvariablen noch empirisch besetzt sind. Angesichts der schon erwähnten Erfordernisse ausreichender Zellenbesetzungen für die

folgende Feinanalyse zeigt eine nähere Inspektion dieser Kreuztabelle, daß die Besetzungen für eine Reihe weiterer Ausprägungskombinationen schon zu schwach sind. Aus diesem Grund scheidet das Merkmal SCHICHT hier aus¹².

Die Partialtafelanalyse erbringt für jede der beiden verbliebenen Kandidatenvariablen keine signifikanten Werte mehr. Damit ist das schrittweise Auswahlverfahren zu einem Ende gekommen. Nach dem Ergebnis dieser Prozedur ist nunmehr die abhängige Variable PM in einer Feinanalyse in Abhängigkeit von den drei unabhängigen Merkmalen SCHULBILDUNG (3), GESCHLECHT (2) und ALTER (3) zu analysieren.

Zuvor jedoch noch eine zusätzliche Bemerkung zum Auswahlverfahren. Es ist anhand der konkreten Zahlen wohl deutlich geworden, daß die Zahl der zu betrachtenden Partialtafeln nach den ersten Schritten stark ansteigt. In unserem Schritt 4 sind bereits 18 Partialtafeln bei einer dreistufigen Kandidaten- und einer dreistufigen Zielvariablen zu betrachten, also $18 \times 4 = 72$ Freiheitsgrade¹³. In Anbetracht dieser hohen Zahl von Freiheitsgraden ist es sehr schwierig, noch signifikante Partialkoeffizienten zu erhalten, jedenfalls wenn man mit Umfragedaten üblichen Umfangs arbeitet¹⁴. Aus diesem Grund haben Landis, Heyman und Koch (1978) eine alternative Kennzahl, die sogenannte "Generalisierte Cochran-Mantel-Haenszel"-Statistik, vorgeschlagen. Diese Maßzahl erfaßt eine durchschnittliche partielle Assoziation. Die zugehörige Zahl der Freiheitsgrade bestimmt sich nach dem Typus einer Partialtabelle, hängt damit also nicht von der Zahl der schon ausgewählten Merkmale ab. Die Berechnung dieser Kennzahl ist arbeitsökonomisch nur durch Integration in ein bestehendes Kreuztabellenprogramm zu bewerkstelligen¹⁵, eine Berechnung per Hand und Tischrechner ist sinnvoll nicht möglich.

Der Nachteil dieser Statistik ist, daß lediglich lokal wirksame Assoziationen (mit anderen Worten: Interaktionswirkungen) damit nur schlecht repräsentiert werden. Da die Analyse mit den ersten drei Einflußfaktoren bereits wesentliche Interaktionswirkungen zutage treten läßt - wie wir im folgenden darlegen werden - erschien es in Anbetracht der Software-Probleme vertretbar, auf eine Berücksichtigung dieses alternativen Abbruchkriteriums zu verzichten.

5. Feinanalyse

Der Ausgangspunkt unserer Feinanalyse ist zusammenfassend in Tabelle 7, S.225 dargestellt. Es handelt sich um eine 18 x 3 Kreuztabelle der Kombination der ausgewählten unabhängigen Merkmale mit der Zielvariablen. In der letzten Spalte ist die relative Häufigkeit der Postmaterialisten in der jeweiligen Teilpopulation, die durch die Ausprägungen der unabhängigen Merkmale definiert wird, dargestellt. Dieser Wert kann auch als Wahrscheinlichkeit dafür interpretiert werden, bei dem die jeweilige Teilpopulation definierenden Typus einen Postmaterialisten zu finden. Diese Wahrscheinlichkeit ist beim Typ Nr.17 - also einer über 60jährigen Frau mittlerer Schulbildung - mit 2,5% am geringsten, beim Typ Nr.3 - einem unter 35jährigen Mann hoher Schulbildung - mit 54,8% am höchsten. Bereits eine bloße visuelle Inspektion dieser Tabelle zeigt, daß die Postmaterialismus-Wahrscheinlichkeit bei jungen und bei hochgebildeten Befragten deutlich höher liegt. Wie anfangs schon angemerkt, ist es eine Aufgabe der Feinanalyse, diese spontan gewonnenen Einsichten statistisch abzusichern - oder auch zu widerlegen¹⁶.

Wir überspringen die Einzelheiten der Feinanalyse (vgl. Anm.5) und wenden uns gleich der Darstellung des Endergebnisses in Schema 1, S.226, zu. In dieser baumartigen Darstellung sind sowohl die Effekte der einzelnen Merkmale als auch die Prädiktorwerte für die jeweiligen Teilpopulationen (Typen) angegeben. Die Effekte für die Materialismus-Wahrscheinlichkeit sind dabei zur besseren Unterscheidung in runde Klammern gesetzt. Bei den Prädiktorwerten am Fuß der Tabelle stehen die Wahrscheinlichkeiten für Postmaterialismus über denen für Materialismus.

Man sieht, daß die Bildungsvariable durchgängig den stärksten Einfluß hat, so liegt z.B. die PM-Wahrscheinlichkeit bei den Typen mit hoher Bildung um 13,6 Prozentpunkte über dem generellen Durchschnittswert. Der Einfluß des Alters ist in seinem Ausmaß stark von der Bildungsstufe abhängig. Je höher die Bildung, desto stärker unterscheiden sich Junge und Alte hinsichtlich der PM-Wahrscheinlichkeit, während eine solche Interaktionswirkung in bezug auf die Materialismus-Wahrscheinlichkeit nicht festzustellen ist. (Der Prädiktorwert für den Mischtyp ergibt sich jeweils als Differenz zu 100). Schließlich ist ein Einfluß des Geschlechts nur unter spezifischen Bedingungen feststellbar, nämlich lediglich in den

Tabelle 7: Häufigkeitsverteilung - Postmaterialismus nach (Alter x Geschlecht x Schulbildung)

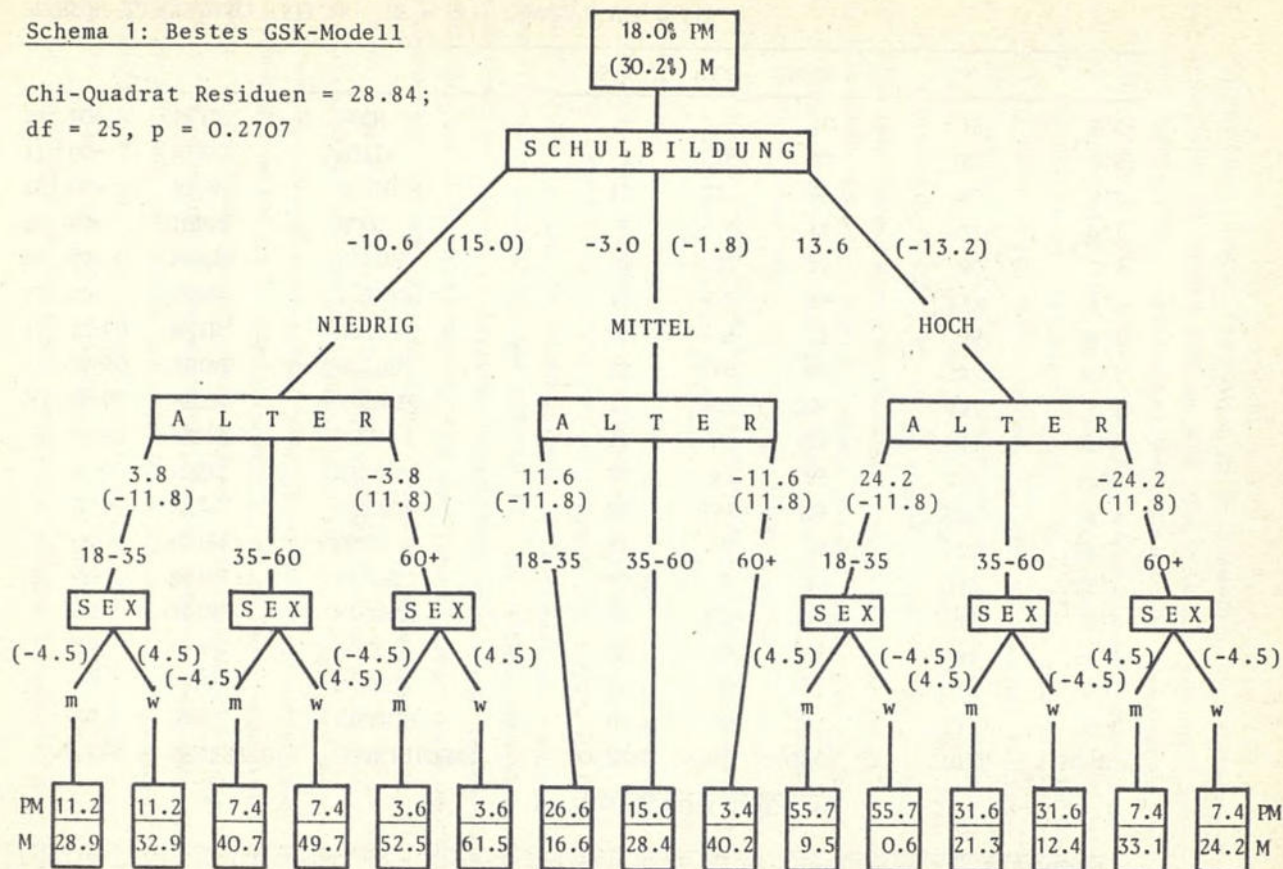
			POSTMATERIALISMUS-TYP			TOTAL	% POSTMAT
ALTER	GESCHLECHT	SCHULBILDUNG	POSTMAT	MAT	MISCH		
1 -35	MÄNNL	NIEDRIG	40	60	117	217	18.4
2 -35	MÄNNL	MITTEL	31	18	70	119	26.0
3 -35	MÄNNL	HOCH	46	10	28	84	54.8
4 -35	WEIBL	NIEDRIG	24	107	114	245	9.8
5 -35	WEIBL	MITTEL	40	29	90	159	25.1
6 -35	WEIBL	HOCH	42	0	35	77	54.3
7 35-60	MÄNNL	NIEDRIG	28	151	206	385	7.3
8 35-60	MÄNNL	MITTEL	26	41	85	152	17.1
9 35-60	MÄNNL	HOCH	14	11	29	54	25.9
10 35-60	WEIBL	NIEDRIG	25	201	206	432	5.8
11 35-60	WEIBL	MITTEL	25	40	88	153	16.3
12 35-60	WEIBL	HOCH	15	8	15	38	39.5
13 60+	MÄNNL	NIEDRIG	12	116	89	217	5.5
14 60+	MÄNNL	MITTEL	3	32	45	80	3.8
15 60+	MÄNNL	HOCH	2	6	14	22	9.1
16 60+	WEIBL	NIEDRIG	13	227	125	365	3.6
17 60+	WEIBL	MITTEL	2	38	40	80	2.5
18 60+	WEIBL	HOCH	1	5	12	18	5.6
			389	1100	1408		

PEARSON CHI-QUADRAT: 571.98; DF = 34; CRAMER'S V = 0.314

Schema 1: Bestes GSK-Modell

Chi-Quadrat Residuen = 28.84;

df = 25, p = 0.2707



Gruppen mit entweder niedriger oder hoher Bildung, und in diesen beiden Gruppen ist er gegenläufig. Frauen niedriger Bildung sind eher materialistisch als Männer dieser Bildungsstufe, während höhergebildete Männer stärker zum Materialismus neigen als entsprechend gebildete Frauen¹⁷.

Zusammenfassend können wir also feststellen, daß sich schwerpunktmäßig die Postmaterialisten relativ eindeutig sozio-demographisch bei den Hochgebildeten der jüngeren Altersjahrgänge lokalisieren lassen. Wenn sich auch keine Differenzierung hinsichtlich des Geschlechts ergibt, so ist zumindest der Mischtypus bei den Frauen häufiger anzutreffen; wenn man so will, das Postmaterialismus-Potential bei den Frauen größer.

Umgekehrt ist die Wahrscheinlichkeit für 'Materialismus' bei den älteren, niedriger Gebildeten am höchsten, wobei die Frauen stärker als die Männer in diese Richtung neigen.

Bezogen auf die eingangs gestellte Untersuchungsfrage legt das Ergebnis dieser Untersuchung nahe, mit "qualitativen" Verfahren arbeitende Fallstudien in den so bestimmten Extremgruppen anzusiedeln.

6. Replikationen

Da auch elaborierte statistische Verfahren nicht davor schützen können, aufgrund einer entweder systematisch verzerrten oder zufällig extrem liegenden Stichprobe zu eher artifiziellen Ergebnissen zu kommen, liegt es nahe, eine Replikation des anhand der Daten des ALLBUS 1980 gefundenen Modells zu versuchen.

In Anbetracht der Beliebtheit, der sich die Inglehartschen Items erfreuen, ist es nicht schwer, geeignete Daten zu finden. Die Replikationen erfolgten anhand der in den Jahren 1978 und 1980 erhobenen Wohlfahrts-surveys¹⁸. Insbesondere die in etwa dem gleichen Zeitraum erhobenen Daten des Wohlfahrtssurvey 1980 zeigen eine in hohem Maße ähnliche Struktur. Eine genauere Übersicht gibt die nachfolgende Tabelle:

Tabelle 8: Vergleich ALLBUS mit Wohlfahrtssurveys

Effekt	ALLBUS '80	RWS '80	RWS '78
mittlerer Anteil PM	18.0%	18.1%	16.2%
mittlerer Anteil M	30.2%	33.1%	35.1%
niedrige Bildung (PM)	-10.6%	-12.3%	-12.4%
hohe Bildung (PM)	13.6%	14.8%	18.2%
niedriges Alter bei			
niedriger Bildung (PM)	3.8%	3.3%	-
mittl. Bildung (PM)	11.6%	9.8%	8.5%
hoher Bildung (PM)	24.2%	24.0%	30.6%
niedrige Bildung (M)	15.0%	15.4%	13.0%
hohe Bildung (M)	-13.2%	-16.1%	-10.1%
niedriges Alter (M)	-11.8%	-14.2%	nach Bildungs- stufen ¹⁹
Geschlecht (PM)	-	2.2%	-
Geschlecht bei			
niedriger Bildung (M)	-4.5%	-8.2%	-4.1%
hoher Bildung (M)	+4.5%	-	-
Chi ² -Residuen	28.84	27.15	29.06
df	25	24	24
P-Wert	0.2707	0.2974	0.2177

Während Alters- und Bildungseinflüsse in den 80er Erhebungen fast identisch sind, ergeben sich hinsichtlich des Geschlechtseinflusses kleinere Verschiebungen. Nach dem Wohlfahrtssurvey 1980 ist ein zwar schwacher, jedoch statistisch signifikanter durchgängiger Einfluß des Geschlechts vorhanden: Männer sind tendenziell eher post-materialistisch orientiert. Hinsichtlich einer geschlechtsspezifischen Materialismusneigung bleibt das Grundmuster gewahrt: Frauen niedriger Bildung sind eher materialistisch orientiert. Jedoch gleicht sich dieser Unterschied mit wachsender Bildung aus (RWS 80) bzw. kehrt sich um (ALLBUS 1980). In Anbetracht der Dif-

ferenziertheit des Modells sind diese Abweichungen von minderer Bedeutung, so daß die Replikation als gelungen anzusehen ist.

Die Übereinstimmung mit den Daten aus dem Jahre 1978 ist etwas schlechter, jedoch bleibt auch hier das prinzipielle Muster erhalten. Die 1978er Daten implizieren insbesondere, den Alterseffekt auch für die Materialismuseigung nach Bildungsstufen getrennt zu erfassen.

7. Ausblick

Wir haben versucht, in diesem Aufsatz Techniken vorzustellen, die unter realistischen Annahmen über die Meßqualität der Daten sowohl formalstatistisch elaboriert verfahren als auch eine Darstellung der Endergebnisse in einer Form ermöglichen, die auch einem bloßen Rezipienten sozialwissenschaftlicher Forschung zugänglich ist. Zugleich versteht sich diese Arbeit als kleiner praktischer Beitrag zu dem Versuch, die unfruchtbare Spaltung in "quantitative" vs. "qualitative" Forschung zu überwinden. Elaborierte statistische Analysen stehen nicht prinzipiell im Widerspruch zu einem "interpretativen" Verständnis von Sozialforschung²⁰; im Gegenteil, mit Bedacht eingesetzt können sie wesentlich helfen, für die Untersuchung in Detailprozessen notwendige 'qualitative' Studien forschungsökonomisch günstig zu platzieren.

Anmerkungen

- 1 Dies verweist natürlich zugleich auf die Notwendigkeit einer theoretischen Präzisierung des Konstrukts. Anderenfalls macht es in einer traditionellen Konzeption von "Gültigkeit" keinen Sinn, diese empirisch überprüfen zu wollen.
- 2 Eine Kluft zwischen den - wenn auch teilweise theoretisch nicht mal sonderlich elaborierten - Konzepten der Sozialforscher und der Praxis der gedanklichen Strukturierung der Umwelterfahrung durch die (befragten) Individuen ist in der politischen Soziologie vielfältig festzustellen. Vgl. dazu etwa Küchler (1982a).
- 3 Dieses Argument ist näher ausgeführt in Küchler (1980 bzw. 1981).
- 4 Eine Ausnahme bildet das von Sonquist und Morgan entwickelte THAID-Verfahren, dessen Automatismus zwar sehr forschungsökonomisch, aber aufgrund seiner jeweils nur lokalen Optimierung nicht in allen Fällen der inhaltlichen Problemstellung angemessen ist.

- 5 Da dieser Ansatz nun auch in der deutschsprachigen Literatur hinreichend dargestellt ist (vgl. u.a. Küchler 1978, 1979, 1982b sowie Küchler und Schwedler 1980), werden wir auf die methodische Darlegung dieses Verfahrens verzichten und lediglich die Ergebnisse präsentieren.
- 6 Damit erfolgt implizit bei nur vier Items ein volles 'ranking'.
- 7 Die im folgenden dargestellten Analysen wurden im Sommer 1980 auf Grundlage eines 'early release' Files des ALLBUS 1980 gerechnet.
- 8 Diese Liste ist hier aus pragmatischen Gründen schon verkürzt worden. Inhaltlich kommen durchaus Merkmale wie Ortsgröße (als allerdings problematischer Indikator für den Stadt-Land-Gegensatz), Familienstand oder ein über den Beruf des Haushaltsvorstands definiertes Schicht-Kriterium (zusätzlich zur hier aufgeführten Selbsteinstufung) in Betracht. Das benutzte schrittweise Auswahlverfahren ist jedoch noch nicht angemessen EDV-technisch umgesetzt, so daß der Arbeits- und Rechenaufwand deutlich mit dem Umfang der Liste der möglichen Erklärungsfaktoren zunimmt.
- 9 Die hier gewählte und allgemein übliche Konstruktion der Typen schöpft die erfragten Informationen (volle Präferenzordnung) nicht aus. Datenanalytisch anspruchsvoller können die $4! = 24$ theoretisch möglichen Präferenzordnungen zunächst unterschieden werden, und eine stärkere Zusammenfassung erst auf Grundlage empirischer Analysen erfolgen (vgl. hierzu Forthofer und Lehen (1981), die diese Strategie anhand der Daten des ALLBUS 1980 demonstrieren). Auch die einfachere Methode der Berücksichtigung der ersten Präferenz bei einer Differenzierung der Mischtypen erweist sich empirisch als gehaltvoll (persönliche Mitteilung von Werner Hagstotz).
- 10 Man beachte, daß im ersten Schritt - der einer ganz konventionellen Kreuztabellenanalyse entspricht - der Einfluß des Geschlechts als völlig vernachlässigbar erschien. Offenbar liegt der Fall einer "Unterdrückung" vor, um im Sprachgebrauch etwa der Lazarsfeldschen Tabellenanalyse zu bleiben (vgl. Zeisel 1970).
- 11 Dieses Kriterium entspricht damit der Signifikanzprüfung in Schritt 1.
- 12 Ähnliche, wenn auch etwas großzügigere Besetzungsvoraussetzungen müssen auch für das schrittweise Auswahlverfahren selbst erfüllt sein. Man könnte jedoch prinzipiell innerhalb der Kombivariablen - und auch in der Feinanalyse - bestimmte Kombinationen zusammenfassen, um diese Besetzungsvoraussetzungen wieder herzustellen. Abgesehen vom Mehraufwand bei der technischen Durchführung verliert die Feinanalyse dadurch jedoch auch an intuitiver Nachvollziehbarkeit.
- 13 In unserem konkreten Fall waren jeweils nur 70 Freiheitsgrade zu berücksichtigen, da jeweils eine der Partialtafeln zu einer 2×3 -Tafel degenerierte und so nur zwei statt vier Freiheitsgrade lieferte.
- 14 Zur anders gelagerten Problematik bei der Analyse von Massendaten vgl. Küchler und Schwedler (1980).
- 15 Für einen geübten Programmierer ist dies mit vertretbarem Aufwand möglich. Das Kölner Zentralarchiv für empirische Sozialforschung verfügt über eine entsprechende lokale Modifikation von SPSS.

- 16 Solche ersten rein intuitiv gewonnenen Einsichten können nämlich auch statistischen Zufallsschwankungen geschuldet sein, deren Kontrolle sich der Intuition entzieht.
- 17 Wir haben, um die Übersichtlichkeit zu wahren, in dieser Darstellung die Angaben der Fehlertoleranzen für die aus dem Modell resultierenden Prädiktorwerte weggelassen. Derartige Werte stehen aber innerhalb des GSK-Ansatzes unmittelbar zur Verfügung. Das NONMET-Programm als EDV-mäßige Umsetzung druckt die Kovarianzmatrix der Prädiktorwerte aus, die Wurzeln aus den Diagonaleinträgen sind damit die Standardabweichungen. Unter Zugrundelegung der Normalapproximation gibt man gewöhnlich den zweifachen Wert der Standardabweichung als Fehlertoleranz (auf dem 5%-Niveau) an.
- 18 Ich danke Prof. Wolfgang Zapf und Dipl.-Soz. Wolfgang Bracht (Sonderforschungsbereich 3 'Mikroanalytische Grundlagen der Gesellschaftspolitik', Universität Mannheim) für die Bereitstellung entsprechender Kreuztabellierungen.
Eine detaillierte Darstellung der aus diesem Material gewonnenen Resultate durch Wolfgang Bracht ist in Vorbereitung.
- 19 Die Alterseffekte belaufen sich hier auf:
-10.1 (niedrig), -7.5 (mittel), -23.0 (hoch).
- 20 Vgl. hierzu auch Kriz (1981) und seine Diskussion von Forschungstechniken als Elemente eines Diskurses.

Literatur

- Barnes, Samuel H. und Max Kaase (Hrsg.), 1979: Political Action. Beverly Hills: Sage.
- Forthofer, Ronald H. und Robert G. Lehnen, 1981: Public Program Analysis: A New Approach to Categorical Data. Practical Statistical Techniques and Methods for Analyzing and Evaluating Public Affairs Programs. Belmont, Cal.: Lifetime Learning (Wadsworth).
- Grizzle, J.D., C.F. Starmer und G.G. Koch, 1969: Analysis of Categorical Data by Linear Models. Biometrics, 25: 489-504.
- Higgins, James E. und Gary G. Koch, 1977: Variable Selection and Generalized Chi-square Analysis of Categorical Data applied to a Large Cross-sectional Occupational Health Survey. International Statistical Review, 45: 51-62.
- Inglehart, Ronald F., 1977: The Silent Revolution. Princeton: Princeton University Press.
- Kriz, Jürgen, 1981: Methodenkritik empirischer Sozialforschung. Eine Problemanalyse sozialwissenschaftlicher Forschungspraxis. Stuttgart: Teubner.
- Küchler, Manfred, 1978: "Alternativen in der Kreuztabellenanalyse: Ein Vergleich zwischen Goodmans 'General Model' (ECTA) und dem Verfahren gewichteten Regression nach Grizzle et al. (NONMET)". Zeitschrift für Soziologie, 7: 347-365.
- Küchler, Manfred, 1979: Multivariate Analyseverfahren. Stuttgart: Teubner.

- Küchler, Manfred, 1980: "Qualitative Sozialforschung: Modetrend oder Neuanfang". Kölner Zeitschrift für Soziologie und Sozialpsychologie, 32: 373-386.
- Küchler, Manfred, 1981: "Der Wahlforscher und die Lebenswelt des Wählers", Politische Vierteljahresschrift, 21, Heft 4.
- Küchler, Manfred, 1982a: "Staats-, Parteien- oder Politikverdrossenheit?". S.39-54, in: J.Raschke (Hrsg.), Bürger und Parteien. Op-laden: Westdeutscher Verlag.
- Küchler, Manfred, 1982b: "Multivariate Analyse nicht-metrischer Daten: Ein Verfahren für die Forschungspraxis". In: H. Kury (Hrsg.), Methodische Probleme kriminologischer Forschung. Köln: Heymanns (i.V.).
- Küchler, Manfred und Erhard Schwedler, 1980: "Die Analyse von kreuz-tabellierten Massendaten: Eine Diskussion neuerer Verfahren" Allgemeines Statistisches Archiv, 64, Heft 4: 360-389.
- Landis, J.R., E.R. Heyman und G.G. Koch, 1978: "Average Partial Association in Three-Way Contingency Tables: A Review and Discussion of Alternative Tests". International Statistical Review, 46: 237-254.
- Zeisel, Hans, 1970: Die Sprache der Zahlen. Köln: Kiepenheuer & Witsch.

SOZIOÖKONOMISCHE DETERMINANTEN VON EINSTELLUNGEN ZU GASTARBEITERN¹

Cornelia Krauth und Rolf Porst

1. Einleitung

Die Bundesrepublik Deutschland verzeichnet, trotz Anwerbestop für Arbeitnehmer aus Nicht-EG-Staaten im November 1973 und deutlichen Bemühungen um restriktivere Asylbestimmungen, sowohl in absoluten als auch in relativen Zahlen die stärkste Ausländerpopulation in ihrer Geschichte. Letzten offiziellen Angaben (Stand: 30.9.1981) zufolge leben derzeit rund 4,6 Millionen Ausländer im Bundesgebiet; damit beträgt ihr Anteil an der Gesamtbevölkerung etwa 7,5% (Statistisches Bundesamt Wiesbaden, 1982: 41).

Der weitaus größte Teil der in der Bundesrepublik lebenden Ausländer kommt aus der Türkei (1,55 Mio. oder 33%); es folgen Jugoslawen (0,64 Mio. oder 14%), Italiener (0,62 Mio. oder 14%) und Griechen (0,30 Mio. oder 7%) (Statistisches Bundesamt Wiesbaden, 1982: 43).

Etwa die Hälfte der in der Bundesrepublik lebenden Ausländer sind Arbeitnehmer (2,07 Mio.) oder registrierte Arbeitslose (97 000), also die in der offiziellen Terminologie als "ausländische Arbeitnehmer" bezeichneten Personen. Ihr Anteil an der Zahl aller Beschäftigten in der Bundesrepublik liegt bei 9,9% und damit nur knapp unter der bisher höchsten "Ausländerquote" (= Anteil der Ausländer an der Gesamtbeschäftigung) von 10,3% im Jahre 1975 (Bundesminister für Arbeit und Sozialordnung 1981).

Diese Zahlen beschreiben den gesellschaftlichen Kontext für die Entwicklung von Einstellungen der deutschen Bevölkerung zu Ausländern: Auf der einen Seite bewirkt allein die große Zahl der Ausländer in der Bundesrepublik ihre verstärkte "Sichtbarkeit" und eine erhöhte Chance für Handlungskonfrontationen. Auf der anderen Seite werden, bei sich verschlechternder wirtschaftlicher Lage und hoher Arbeitslosigkeit in der Bundesrepublik, Ausländer verstärkt als Konkurrenten auf dem Arbeitsmarkt oder als Wirtschaftskonkurrenten allgemein wahrgenommen.

Diese Entwicklung bleibt nicht ohne Einfluß auf die Einstellungen der bundesdeutschen Bevölkerung zu Gastarbeitern. Die Veränderung des Meinungsklimas ist bereits Gegenstand der Diskussion in den Massenmedien. So befaßte sich das Nachrichtenmagazin "DER SPIEGEL" in seiner Ausgabe vom 15. September 1980 in der Titelgeschichte mit dem "Fremdenhaß in der Bundesrepublik": Eine "neue Welle von Ausländerfeindlichkeit" sei zu verzeichnen, und wenngleich sicher sei, daß massivere Formen von Aggression gegen Ausländer nach wie vor vereinzelte Phänomene und vereinzelt extremistischen Tätern anzulasten seien, so spreche doch manches dafür, "daß die Gewalt nur der sichtbare Ausdruck einer Stimmung ist, die sich im Bundesvolk breitgemacht hat".

2. Problemstellung

Wir wollen im folgenden der Frage nachgehen, ob die Abneigung gegen Ausländer in der Tat als repräsentatives Einstellungsmuster in der bundesdeutschen Gesellschaft vorhanden ist, oder ob die Einstellungen zu Ausländern gerade wegen ihrer Wahrnehmung als Wirtschaftskonkurrenten nicht viel eher abhängig sind von im weiteren Sinne sozioökonomischen Merkmalen der Einstellungsträger.

Mit Hilfe eines multivariaten Analyseverfahrens (LISREL) wollen wir deshalb nach Zusammenhängen fragen zwischen dem objektiven und subjektiven Status von Personen und ihren Einstellungen zu Gastarbeitern. Weiterhin wollen wir prüfen, ob diese Einstellungen durch tatsächliche Kontakte zu Gastarbeitern beeinflußt werden.

Durch LISREL-Gruppenvergleiche wollen wir nachweisen, daß auch die Dimension der Wettbewerbserfahrung bzw. Wettbewerbserwartung im sozioökonomischen Bereich, vor allem die Konkurrenz um Arbeitsplätze, zu einer stärkeren Diskriminierungsbereitschaft gegenüber Gastarbeitern führt.

Wir gehen von der Annahme aus, daß der Wahrnehmung von Ausländern als sozioökonomischer Konkurrenz, insbesondere als Konkurrenz auf dem Arbeitsmarkt, zentrale Bedeutung zukommt und unterstellen, daß eine solche Konkurrenz um so deutlicher wahrgenommen wird, je geringer der

objektive und subjektive Status der Befragten ist. Wir behaupten weiterhin, daß Personen, die arbeitslos sind oder waren oder Arbeitslosigkeit befürchten, negativere Einstellungen zu Ausländern haben als Personen ohne solche Erfahrungen bzw. Befürchtungen, weil auch sie Gastarbeiter eher als Konkurrenten auf dem Arbeitsmarkt wahrnehmen.

Als Datenbasis benutzen wir die Ergebnisse des ALLBUS 1980, einer für die wahlberechtigte Bevölkerung der Bundesrepublik und West-Berlins repräsentativen Mehrthemen-Querschnittserhebung mit 2955 Befragten, die Anfang 1980 durchgeführt worden ist.

3. Gastarbeiter als Minderheiten

Die Bezeichnung "Gastarbeiter" hat sich in zunehmendem Maße als unbrauchbar erwiesen zur Kategorisierung der Ausländerpopulation in der Bundesrepublik. Hatte der Begriff zunächst noch einen deskriptiven Sinn insofern, als Ausländer tatsächlich als Arbeitskräfte (und in der Tat zu meist als Arbeiter) auf Zeit hier lebten (oder leben sollten), so hält sich heute etwa die Hälfte der Ausländer explizit nicht (oder noch nicht) als Arbeitnehmer in der Bundesrepublik auf. Zugleich betrachten viele Ausländer ihren Aufenthalt hier nicht mehr nur als kurzzeitig befristet, sondern streben einen dauerhaften, zumindest aber langfristigen Verbleib an.

Der Begriff "Gastarbeiter" hat heute einen weitgehend negativ besetzten Bedeutungsgehalt: Ausländer werden dann als "Gastarbeiter" klassifiziert, wenn ihre "Fremdartigkeit" aufgrund bestimmter äußerlicher Merkmale wahrgenommen wird und wenn sie zugleich direkt oder indirekt (Familienangehörige) als tatsächliche oder potentielle Inhaber von Berufsrollen auf dem deutschen Arbeitsmarkt auftreten, ohne dabei einen überdurchschnittlich hohen Berufsstatus zu besitzen. Unter diese Definition fallen vor allem Personen aus süd- und südosteuropäischen sowie asiatischen Ländern, die in der Bundesrepublik in der Regel in sozial niedrig bewerteten und in den Arbeitsbedingungen entsprechend schlecht ausgestatteten Berufspositionen tätig sind, sowie deren Familien.

Gastarbeiter können als Minderheit im Sinne der Definition von Rose (1969/1972) bezeichnet werden, sind also zu beschreiben als "besondere

Rassen-, Kultur-, Religions- oder Nationalitätengruppen, die, inmitten anderer Gruppen lebend, doch nicht voll teilhaben an der allgemeinen Kultur, der sie als Teil angehören".

Als Minderheiten werden ihnen typischerweise gewisse ökonomische, soziale und politische Rechte vorenthalten, d.h. sie unterliegen typischerweise der Diskriminierung durch die Mehrheit: Wenn die Mitglieder einer Gesellschaft in der Regel mit bestimmten (kodifizierten oder gewohnheitsmäßig zugebilligten) politischen, sozialen und gesellschaftlichen Rechten ausgestattet sind, so verstehen wir unter Diskriminierung Verhalten, im speziellen Falle hier auch verbales Verhalten, das darauf ausgerichtet ist, gewisse Teilgruppen der Gesellschaft von der Teilhabe an diesen Rechten auszuschließen bzw. ihnen diese Teilhabe abzusprechen. Beruhen solche Verhaltensweisen und Versuche auf ethnischer Kategorisierung, wird also Mitgliedern ethnischer Minderheiten durch die Mehrheit aufgrund ethnischer Unterschiede und Unterscheidungen eine gleichberechtigte Teilhabe an solchen Rechten vorenthalten bzw. abgesprochen, sprechen wir von ethnischer Diskriminierung. Diskriminierung kann als Ergebnis eines Prozesses der Realisierung von Vorurteilen verstanden werden (Schäfer und Six 1978: 224).

Diskriminierung als Verhalten oder Verhaltensbereitschaft richtet sich allgemein gegen der Mehrheit fremde oder von ihr als fremd definierte Verhaltensweisen und Erscheinungsformen und spezifisch gegen eine gleichberechtigte Existenz von Minderheiten in unterschiedlichen sozialen Handlungsfeldern. Als Diskriminierung eher allgemeiner und unspezifischer Art läßt sich die Forderung interpretieren, Gastarbeiter sollten ihren Lebensstil an denjenigen der deutschen Gastgeber anpassen.

Anpassung des Lebensstils: Die Chance der Integration der Gastarbeiter in die Gesellschaft der Gastgeber (im Sinne einer "Teilhabe an gesellschaftlichen Werten" wie z.B. berufliche Stellung und Einkommen sowie an gesellschaftlichen Organisationen) ist von direktem Einfluß auf den Grad ihrer Assimilation (im Sinne einer "Teilhabe an Kulturelementen" wie Sprache, Werte, Normen) (Hoffmann-Nowotny 1976: 57 und 61). Hoffmann-Nowotny (1976: 57) folgert aus dieser, seiner Ansicht nach gut bestätigten Hypothese: "Wenn also von den Gastarbeitern verlangt wird, sie sollten sich als erstes den Bräuchen, Gewohnheiten und Sitten des Ein-

wanderungslandes anpassen ..., so ist dagegen aus soziologischer Sicht einzuwenden, daß Assimilationsbereitschaft und Assimilation der Gastarbeiter primär eine Funktion der Integrationsbereitschaft des aufnehmenden Landes sind."

Wenn aber Gastarbeitern die Teilhabe an bestimmten politischen, sozialen und gesellschaftlichen Rechten als Integrationsvoraussetzung ganz oder teilweise vorenthalten wird, ist nicht zu erwarten, "daß sie besondere Anstrengungen zur Teilhabe an der Kultur des Einwanderungslandes auf sich nehmen" (Hoffmann-Nowotny 1976: 57). Eine wenngleich abgeschwächt formulierte Forderung nach Anpassung des Lebensstils, also nach Verzicht auf gewisse eigene gesellschaftliche und kulturelle Traditionen und damit tendenziell nach Aufgabe der eigenen kulturellen Identität bei gleichzeitiger Behinderung der Integration in die Gesamtgesellschaft stellt eine Diskriminierung in dem oben definierten Sinne dar.²

Neben der eher allgemein gehaltenen Forderung nach Anpassung der Gastarbeiter an den Lebensstil der deutschen Bevölkerung äußert sich Diskriminierung auch in relativ deutlich abgrenzbaren sozialen Handlungsfeldern, z.B. im Arbeitsbereich, in der Ausübung politischer Rechte oder im privaten Bereich.

Arbeitsbereich: Die Bedeutung der Gastarbeiter für die wirtschaftliche Entwicklung in der Bundesrepublik und ihr Einfluß auf den deutschen Arbeitsmarkt sind politisch (s. etwa Bundesminister für innerdeutsche Beziehungen, Hrsg., 1974: 16ff. oder Nikolinakos 1974), journalistisch (s. etwa DER SPIEGEL, 29, 1975, Hefte 3, 18 und 50) und wissenschaftlich (s. etwa Kühl, 1976) ausgiebig diskutiert.

Seit geraumer Zeit hat sich diese Diskussion verlagert in Richtung auf die Frage, ob eine teilweise oder völlige "Rückführung" der Gastarbeiter in ihre Heimatländer nicht ein geeignetes Mittel zur Lösung gesellschaftlicher Probleme in der Bundesrepublik, insbesondere zur Senkung der Arbeitslosenzahl, sein könnte. Unabhängig davon, wie begründet die Vermutung eines Zusammenhangs zwischen der Beschäftigung ausländischer Arbeitnehmer und der Arbeitslosenquote in der Bundesrepublik ist, stellt die Forderung, man solle bei zunehmender Knappheit von Arbeitsplätzen Gastarbeiter wieder in ihre Heimat zurückschicken, eine massive Diskriminierung dieser Bevölkerungsgruppen dar. Die Definition ihrer Lebensum-

stände wird dadurch reduziert auf eine disponible, ausschließlich den Bedürfnissen des Gastgeberlandes verpflichtete Ausübung zugewiesener segmentärer Rollen. Gastarbeiter wären dann in der Tat eine "industrielle Reservearmee" (vgl. Katsoulis 1978: 17ff.), die je nach Bedarf dem deutschen Arbeitsmarkt zugeführt und wieder entzogen werden könnte. Eine solche Vorstellung über die Verfügbarkeit von Gastarbeitern läuft konträr mit Bemühungen um ihre Integration bzw. schließt Bemühungen um ihre Integration von vornherein aus.

Wenn aber Integration die möglichst weitgehende und gleichberechtigte Teilhabe an den gesellschaftlichen Prozessen des Gastgeberlandes bedeutet, führt die ausschließlich instrumentelle Definition der Gastarbeiter als Träger von Berufsrollen, deren Dauer und Ausübung durch das Gastgeberland reglementiert sind, automatisch zum Ausschluß von der Teilhabe an bestimmten sozialen und gesellschaftlichen Rechten. Wenn darüber hinaus der Aufenthalt in der Bundesrepublik abhängig gemacht wird von durch den einzelnen nicht kontrollierbaren und nicht beeinflussbaren sozioökonomischen Entwicklungen wie z.B. der Verknappung von Arbeitsplätzen, so verringert sich für Gastarbeiter in erheblichem Maße die Chance einer individuellen Daseinsplanung, die etwa im Sinne eines Rechts auf freie Entfaltung der Persönlichkeit jeder auf deutschem Staatsgebiet lebenden Person durch das Grundgesetz für die Bundesrepublik Deutschland gewährleistet werden soll.

Politische Betätigung: Der Gesetzgeber schließt nach wie vor Ausländer, die in der Bundesrepublik leben, von der aktiven und passiven Teilnahme an Wahlen auf Bundes-, Landes- und Kommunalebene aus. Allerdings ist die Teilnahme von Ausländern an Kommunalwahlen immerhin Gegenstand der politischen Diskussion.³

Wir brauchen uns an dieser Stelle mit dem Problem des Ausländerwahlrechts materiell nicht zu befassen. Wenn aber das Wahlrecht ein fundamentaler und zentraler Bestandteil demokratischer Partizipation ist und wenn andererseits Diskriminierung auf die Vorenthaltung von Rechten hinzielt, so muß die Vorenthaltung des Wahlrechts als massive Diskriminierung der Gastarbeiter interpretiert werden.

Sieht man vom Wahlrecht ab, gesteht der Gesetzgeber ausländischen Personen in der Bundesrepublik grundsätzlich das Recht auf politische Be-

tätigung zu. Formal kommen sie in den Genuß aller Grundrechte, welche das Grundgesetz nicht ausdrücklich deutschen Staatsbürgern vorbehält. So billigt das Grundgesetz zum Beispiel jedem, damit auch den in der Bundesrepublik lebenden Gastarbeitern, das Recht auf freie Meinungsäußerung zu. Aufgrund einfacher Bundesgesetze können Ausländer das Recht auf Versammlungs- und Vereinigungsfreiheit in Anspruch nehmen, welches das Grundgesetz an sich auf Deutsche beschränkt, und sie genießen das politische Demonstrationsrecht (vgl. hierzu Katsoulis 1978: 99ff. oder Tomuschat 1974). Allerdings darf man nicht übersehen, daß das Ausländergesetz deutsche Behörden mit weitgehenden Möglichkeiten zur Einschränkung dieser Rechte ausstattet (Beispiele bei Katsoulis 1978: 102ff.).

Die verbreitete Forderung, man solle Gastarbeitern jede politische Betätigung in der Bundesrepublik untersagen, stellt insofern eine Diskriminierung der Gastarbeiter dar, als sie darauf abzielt, ihnen die Teilhabe an politischen Rechten streitig zu machen, die nicht nur den Deutschen selbstverständlicherweise zugestanden werden, sondern per Gesetz sogar den in der Bundesrepublik lebenden Ausländern selbst.

Privatbereich: Bei 7,7% der 1980 in der Bundesrepublik geschlossenen Ehen war einer der Partner Ausländer (Statistisches Bundesamt, Wirtschaft und Statistik 1/82: 38). Gegenüber 1950 mit 3,6% und 1960 mit 3,7% (Presse- und Informationsamt der Bundesregierung, Hrsg., Gesellschaftliche Daten 1979: 25) hat sich die ethnische Mischehe somit zu einem demographisch relevanten sozialen Phänomen entwickelt.

Wenn gefordert wird, Gastarbeiter sollten ihre Ehepartner unter ihren eigenen Landsleuten aussuchen, so stellt dies einen Versuch der Reglementierung der Privatsphäre der Gastarbeiter dar, der explizit durch nichts anderes begründet wird als durch ethnische Kategorisierung. Die Diskriminierung besteht nicht darin, daß Gastarbeiter in der Tat unter sich bleiben (müssen), also ihre Ehepartner tatsächlich weitgehend unter ihren eigenen Landsleuten auswählen, sondern darin, daß die Majorität es von ihnen fordert. D.h. die Majorität stellt eine Forderung, deren Realisierung sogar oft im Interesse der Gastarbeiter selbst liegen könnte und diskriminiert die Gastarbeiter zugleich, weil sie diese Forderung stellt (oder im Sinne von Esser 1980: 145: Beteiligungsausschlüsse zuschreibt).

4. Ansätze zur Erklärung von Einstellungen zu Gastarbeitern

Zur Erklärung der Diskriminierung von Minderheiten bieten sich zwei Ansätze an. Ein sozialpsychologischer Ansatz, auf den wir im folgenden nicht näher eingehen werden, gründet auf der Annahme, daß Gruppen zur Wahrung ihrer Integration Fremdgruppen definieren, denen sie mit Ablehnung gegenüberstehen. Eine tatsächliche oder vermeintliche Gefährdung der Eigengruppe führt zu einer verstärkten Bereitschaft, die Mitglieder der Fremdgruppe zu diskriminieren, weil diese für die drohende Desintegration der Eigengruppe verantwortlich gemacht werden (Sündenbock-Theorie, vgl. Allport 1954).

Wir konzentrieren uns im folgenden auf einen soziologischen Ansatz, der von der Annahme ausgeht, daß Diskriminierung als Folge von Wettbewerb oder Wettbewerbserwartung auftritt: "Wettbewerbsbefürchtungen und Statusbedrohungen verstärken askriptive Abwehrbereitschaften" (Esser 1980: 146). Mangel an Wettbewerbsfähigkeit wird durch Diskriminierungen auszugleichen gesucht.

Diskriminierungsbereitschaft und Diskriminierungsverhalten sind in diesem Ansatz nicht generalisiert, sondern abhängig von objektiven Wettbewerbssituationen oder von subjektiv wahrgenommenem Wettbewerb. Anders ausgedrückt: Diskriminierungen sollten verstärkt zwischen solchen Personen auftreten, die tatsächlich oder vermeintlich in Konkurrenz zueinander stehen.

Die Wahrnehmung von Gastarbeitern als Konkurrenten konzentriert sich im wesentlichen auf die sozioökonomische Dimension des Wettbewerbs um Arbeitsplätze. Da Gastarbeiter (wenn überhaupt) als Konkurrenten um objektiv niedrig platzierte Berufspositionen agieren, dürfte eine Wettbewerbssituation bzw. Wettbewerbserwartung verstärkt bei solchen einheimischen Personen aufzufinden sein, die aufgrund eigener sozioökonomischer Defizite selbst auf die Ausübung niedrig bewerteter Berufstätigkeiten angewiesen sind. D.h. die Diskriminierung von Gastarbeitern müßte vor allem abhängig sein von der Wettbewerbserfahrung bzw. -erwartung einerseits, von den Voraussetzungen für eine Wettbewerbsbehauptung der Einstellungsträger andererseits.

Die Chancen zur Wettbewerbsbehauptung im sozioökonomischen Bereich, speziell auch auf dem Arbeitsmarkt, sind aber abhängig von statusbildenden Variablen wie der Schulbildung, der Berufsausbildung und der beruflichen Stellung. Neben dem objektiven Status müßte die Wahrnehmung der eigenen Position in der Gesellschaft, also ein subjektiver Status, insofern eine Rolle spielen, als sie die Wahrnehmung individueller Wettbewerbschancen beeinflußt. Das hieße also, Einstellungen zu Gastarbeitern seien, im Zusammenhang mit Wettbewerbserfahrung bzw. -erwartung, determiniert durch Merkmale des objektiven und subjektiven Status der Einstellungsträger.

Eine ausschließliche Konzentration auf Statusmerkmale als Determinanten von Einstellungen zu Gastarbeitern vernachlässigt allerdings die Bedeutung von Kontakten zwischen Mitgliedern von Mehrheits- und Minderheitengruppen als Regulativ für Diskriminierungsbereitschaft. Ob tatsächliche Kontakte Vorurteile und Diskriminierungen vermindern oder im Gegenteil sogar erhöhen, hängt ab von der Art, Intensität, Wichtigkeit und einer Reihe anderer Kriterien der stattfindenden Kontakte (Amir 1969). Beliebige Kontakte alleine lassen nicht von vornherein eine Reduzierung von Vorurteilen und Diskriminierungsbereitschaft erwarten (Allport 1954: 269; Schäfer und Six 1978: 278), sondern schaffen nur Voraussetzungen für Einstellungsänderungen. Wettbewerb ist nach Amir (1969: 338) eine von mehreren ungünstigen Bedingungen, unter denen Kontakte zu einer Verstärkung von Vorurteilen führen können. Bei Personen mit Wettbewerbserfahrung müßten demzufolge tatsächliche Kontakte, wenn überhaupt, dann nur in geringerem Maße zu einer Reduzierung von Diskriminierungsbereitschaft geführt haben als bei Personen ohne solche Erfahrungen, das heißt aber auch, ihre Einstellungen zu Gastarbeitern müßten durchweg weniger positiv sein als bei Personen, die Gastarbeiter nicht als Konkurrenten wahrnehmen.

5. Spezifikation der Variablen

Bevor wir zur Darstellung unseres multivariaten (LISREL-)Modells kommen, wollen wir zunächst die Variablen spezifizieren, welche in das Modell aufgenommen bzw. zur Definition von Teilstichproben für Subgruppenvergleiche verwendet werden sollen.

Wir unterscheiden drei Gruppen von Variablen, nämlich "sozioökonomische" Variablen, Kontakt- und Einstellungs-Variablen. In Klammern geben wir jeweils die Nummer der betreffenden Fragen im Fragebogen des ALLBUS 1980 an.

5.1 "Sozioökonomische" Variablen

Von den sozioökonomischen Variablen gehen in das Modell selbst ein die Variablen

- Schulabschluß des Befragten (S 3)
- Beruflicher Ausbildungsabschluß des Befragten (S 4)
- Netto-Einkommen des Befragten (S 39 G)
- Berufsprestige des Befragten für gegenwärtigen Beruf (synthetische Variable nach Treiman) 4
- Berufsprestige des Befragten für letzten Beruf (synthetische Variable nach Treiman)
- Subjektive Schichteinstufung (22)
- Einstufung auf der Oben-Unten-Skala (S 46)
- Equity = Wahrnehmung sozialer Gerechtigkeit (23).

Zur Definition von Teilstichproben für Subgruppenvergleiche verwenden wir die Variablen

- Aktuelle Arbeitslosigkeit des Befragten (S 5)
- Aktuelle Arbeitslosigkeit des Ehepartners (S 23)
- Furcht vor Arbeitslosigkeit beim Befragten (S 8a)
- Furcht vor Arbeitslosigkeit beim Ehepartner (S 29)
- Erfahrung von Arbeitslosigkeit beim Befragten (S 9)
- Erfahrung von Arbeitslosigkeit beim Ehepartner (S 24).

Diese Variablen korrespondieren im wesentlichen mit der Wettbewerbsdimension der Beziehungen zwischen Deutschen und Gastarbeitern.

5.2 Kontakt-Variablen

Kontakte zu Gastarbeitern (Frage 26) werden erfragt

- in der eigenen Familie oder näheren Verwandtschaft (26a)
- am Arbeitsplatz (26b)
- in der Nachbarschaft (26c) und
- im sonstigen Freundes- und Bekanntenkreis (26d).

Die Kontakt-Variablen messen tatsächlichen Kontakt der Befragten zu Gastarbeitern in unterschiedlichen, sehr wahrscheinlich emotional ungleich

besetzten Kontaktfeldern. Es wird dabei nur die Tatsache eines Kontaktes überhaupt registriert, nicht die Häufigkeit und schon gar nicht die Intensität oder Bewertung (z.B. der Wichtigkeit) der Kontakte durch die Befragten.

Es erscheint uns deshalb ausreichend und den Daten angemessen, die vier Kontakt-Variablen zu einem "Index Kontakte" zusammenzufassen. Wir addieren zu diesem Zweck gleichwertig die Kontakte in den verschiedenen Kontaktfeldern pro Person, so daß der Index Ausprägungen von 0 = "überhaupt kein Kontakt", 1 = "Kontakt in einem Kontaktfeld", etc. bis zu 4 = "Kontakt in allen vier Kontaktfeldern" hat.

5.3 Einstellungs-Variablen

Mit den Einstellungs-Variablen werden diskriminierende Einstellungen gegenüber Gastarbeitern gemessen. Eines der Items bezieht sich dabei auf die eher allgemeine Forderung nach Anpassung des Lebensstils, die anderen drei auf Diskriminierungen in abgegrenzten sozialen Handlungsfeldern, nämlich im Arbeitsbereich, im Bereich der politischen Betätigung und im privaten Bereich (Frage 25; Itemformulierungen im folgenden Abschnitt).

6. Deskriptive Ergebnisse

Wir haben zu Anfang gefragt, ob die Abneigung gegen Ausländer ein repräsentatives Einstellungsmuster der bundesdeutschen Bevölkerung sei. Wäre dies der Fall, müßte bereits aus den Randverteilungen der Einstellungssitems die negative Bewertung der Gastarbeiter zu erkennen sein, d.h. die Einstellungssitems, welche wir als diskriminierende Einstellungen definiert haben, müßten in deutlichem Ausmaß auf Zustimmung der repräsentativen Stichprobe treffen.

Dies gilt sicher für die Forderung nach Anpassung an den Lebensstil der Deutschen, das allerdings zweifelsohne im Sinne einer Diskriminierung am unverbindlichsten formulierte Item: Etwa 65% der Befragten stimmen dieser Forderung zu (d.h. wählen eine Kategorie über dem mittleren Skalenwert), allein fast ein Drittel stimmen sogar "voll und ganz" zu:

Tabelle 1: Anpassung an den Lebensstil

"Gastarbeiter sollten ihren Lebensstil ein bißchen besser an den der Deutschen anpassen."

Skalenwert	Skalenendpunkte	abs.	v.H.
1	stimme überhaupt nicht zu	224	7.6
2		155	5.3
3		235	8.0
4		401	13.6
5		606	20.5
6	stimme voll und ganz zu	444	15.0
7		878	29.7
	verweigert	4	.1
	weiß nicht	2	.1
	keine Angabe	6	.2
		<hr/> 2955	<hr/> 100.1

Weniger Befragte als oben, nämlich rund die Hälfte sind der Ansicht, man solle Gastarbeiter wieder in ihre Heimat schicken, wenn Arbeitsplätze knapp werden und man solle Gastarbeitern jegliche politische Betätigung in der Bundesrepublik untersagen. Ein Viertel der Befragten stimmen der ersten Aussage "voll und ganz" zu, fast ein Drittel der zweiten Aussage. Der hohe Anteil an Personen, die hier diskriminierende Einstellungen verbalisieren, verdient unserer Ansicht nach angesichts der Wichtigkeit der Bereiche Arbeitsleben und politische Betätigung für die individuelle Daseinsgestaltung wie für die kollektive Integration der Gastarbeiter besondere Beachtung.

Tabelle 2: Ausweisung bei Arbeitsplatzknappheit

"Wenn Arbeitsplätze knapp werden, sollte man die Gastarbeiter wieder in ihre Heimat zurückschicken."

Skalenwert	Skalenendpunkte	abs.	v.H.
1	stimme überhaupt nicht zu	470	15.9
2		234	7.9
3		286	9.7
4		417	14.1
5		431	14.6
6		345	11.7
7	stimme voll und ganz zu	756	25.6
	verweigert	4	.1
	weiß nicht	2	.1
	keine Angabe	10	.3
		<hr/> 2955	<hr/> 100.0

Tabelle 3: Verbot politischer Betätigung

"Man sollte Gastarbeitern jede politische Betätigung in Deutschland untersagen."

Skalenwert	Skalenendpunkte	abs.	v.H.
1	stimme überhaupt nicht zu	519	17.6
2		255	8.6
3		291	9.9
4		378	12.8
5		354	12.0
6		255	8.6
7	stimme voll und ganz zu	891	30.2
	verweigert	4	.1
	weiß nicht	4	.1
	keine Angabe	4	.1
		<hr/> 2955	<hr/> 100.0

Am liberalsten erweisen sich die Befragten bei der Beurteilung des Items aus dem Privatbereich, nämlich Partnerwahl der Gastarbeiter. Hier stimmen immerhin ein Viertel "überhaupt nicht" zu, wenn gefordert wird,

Gastarbeiter sollten sich ihre Ehepartner unter ihren Landsleuten auswählen. Zum erstenmal ist es weniger als die Hälfte der Befragten, die diskriminierend antwortet, also diese Forderung unterstützt, wobei ein Viertel "voll und ganz" zustimmt:

Tabelle 4: Wahl des Ehepartners

"Gastarbeiter sollten sich ihre Ehepartner unter ihren Landsleuten auswählen."			
Skalenwert	Skalenendpunkte	abs.	v.H.
1	stimme überhaupt nicht zu	731	24.7
2		240	8.1
3		254	8.6
4		413	14.0
5		328	11.1
6	stimme voll und ganz zu	284	9.6
7		692	23.4
	verweigert	5	.2
	weiß nicht	2	.1
	keine Angabe	6	.2
		<hr/> 2955	<hr/> 100.0

Aus diesen Ergebnissen sind erste Schlüsse über die Diskriminierungsbereitschaft der Befragten zu ziehen. Zwar wird man wohl nicht von durchweg negativen Einstellungen zu Gastarbeitern als repräsentativem Einstellungsmuster sprechen können, doch wird immerhin bei etwa der Hälfte der Befragten eine Diskriminierungsbereitschaft deutlich. Deshalb werden wir in der folgenden multivariaten Analyse nach Merkmalen von Trägern negativer Einstellungen zu Gastarbeitern fragen, und wir werden diese Form von Diskriminierungsbereitschaft im Zusammenhang mit dem objektiven und subjektiven Status der Befragten und ihren Kontakten zu Gastarbeitern untersuchen.

7. Darstellung des Modells

Der erste Schritt der Analyse besteht in der Konstruktion und Darstellung eines LISREL-Modells "Determinanten von Einstellungen zu Gastarbeitern". Die von uns als zentral behauptete und über Erfahrung mit Arbeitslosigkeit und Furcht vor Arbeitslosigkeit operationalisierte Wettbewerbsdimension kann dabei noch vernachlässigt werden. Das folgende Modell 0 soll die von uns erwarteten Zusammenhänge zwischen statusbildenden Merkmalen der Befragten sowie ihren Kontakten zu Gastarbeitern einerseits und ihren Einstellungen zu Gastarbeitern andererseits verdeutlichen.⁵

Für das Modell gilt:

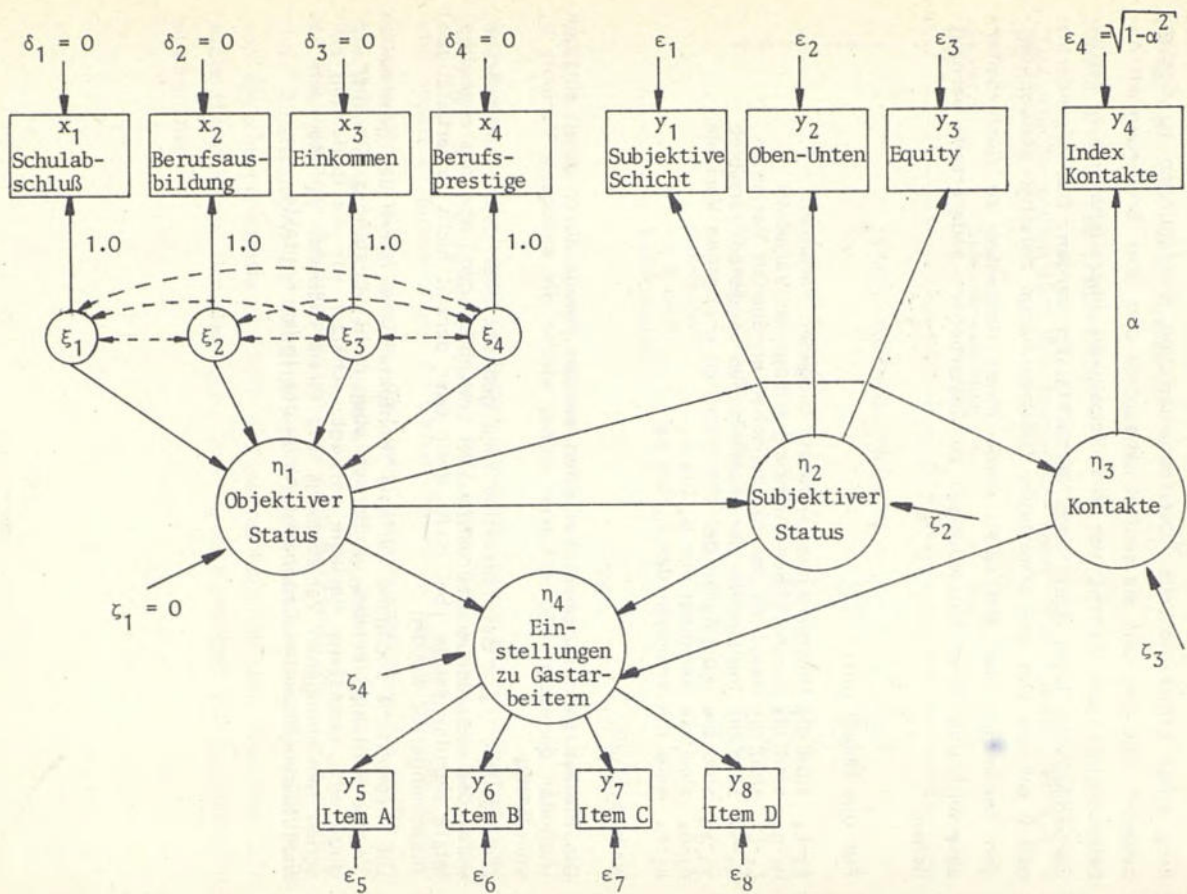
- $\xi_1 - \xi_4$ sind die theoretischen (wahren) exogenen Variablen
- $\eta_1 - \eta_4$ sind die theoretischen (wahren) endogenen Variablen
- $\zeta_1 - \zeta_4$ sind die Residuen der theoretischen endogenen Variablen
- $x_1 - x_4$ sind die Indikatoren der theoretischen exogenen Variablen
- $y_1 - y_8$ sind die Indikatoren der theoretischen endogenen Variablen
- $\delta_1 - \delta_4$ sind die Residuen der x_1 bis x_4
- $\varepsilon_1 - \varepsilon_8$ sind die Residuen der y_1 bis y_8

Weiterhin gilt:

Die theoretischen exogenen Variablen werden jeweils durch einen einzigen Indikator gemessen und erklären selbst wieder die endogene Variable η_1 vollständig.

Die Variable x_4 = Berufsprestige wird gebildet über den Treiman-Score, wobei das aktuelle Berufsprestige (bei Erwerbstätigen) und das ehemalige, letzte Berufsprestige (bei nicht mehr oder derzeit nicht Erwerbstätigen) zusammengefaßt werden.

Die theoretische endogene Variable "Kontakte" wird durch den gemessenen "Index Kontakte" erklärt, sicherlich aber nicht vollständig. Da aber nur diese eine gemessene Variable vorliegt, schätzen wir die Beziehung zwischen der endogenen Variablen η_3 und ihrem Indikator y_4 über den Reliabilitätskoeffizienten Cronbachs Alpha des Index Kontakte.



Modell 0: Determinanten von Einstellungen zu Gastarbeitern

Die Strukturgleichungen für dieses Modell lauten:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ \beta_{21} & 1 & 0 & 0 \\ \beta_{31} & 0 & 1 & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & 1 \end{bmatrix} \cdot \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} 0 & \lambda_{12} & 0 & 0 \\ 0 & \lambda_{22} & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \lambda_{54} \\ 0 & 0 & 0 & \lambda_{64} \\ 0 & 0 & 0 & \lambda_{74} \\ 0 & 0 & 0 & \lambda_{84} \end{bmatrix} \cdot \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \sqrt{1-\alpha^2} \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix}$$

Das Modell ist konventionell insofern, als es die latenten Statusvariablen mit Indikatoren bildet, welche üblicherweise zur Bestimmung des objektiven und subjektiven Status herangezogen werden. Dies gilt vor allem für den "objektiven Status", der sich zusammensetzt aus den gemessenen Variablen "Schulabschluß", "Berufsausbildung", "Einkommen" und "Berufsprestige". Das Modell ist unkonventionell insofern, als wir bewußt davon abgesehen haben, "objektiven Status" als Index unter der Annahme gleicher Beiträge der einzelnen gemessenen Variablen oder a priori behaupteter unterschiedlicher Beiträge dieser Variablen zu konstruieren. Anstelle einer Index-Bildung mit ungeprüften Annahmen über die Zusammensetzung dieser latenten Variablen und über die Beiträge der einzelnen gemessenen Variablen wollen wir "objektiven Status", den Daten folgend, empirisch bestimmen.

Wir prüfen nun das Modell zuerst an der Gesamtstichprobe (Modell 1), danach im Vergleich zwischen verschiedenen Subgruppen in einer simultanen Analyse. Die jeweils direkt miteinander zu vergleichenden Gruppen sind disjunkt:

Modell 2a: Personen mit Erfahrung von Arbeitslosigkeit, also Personen, die selbst bzw. deren Ehepartner arbeitslos sind oder waren ("Arbeitslose")

vs.

Modell 2b: Personen ohne solche Erfahrungen ("Nicht-Arbeitslose")⁶,

Modell 3a: Personen mit Furcht vor Arbeitslosigkeit bei sich selbst oder bei ihrem Ehepartner ("Furcht Arbeitslosigkeit")

vs.

Modell 3b: Personen ohne solche Furcht ("Keine Furcht Arbeitslosigkeit").

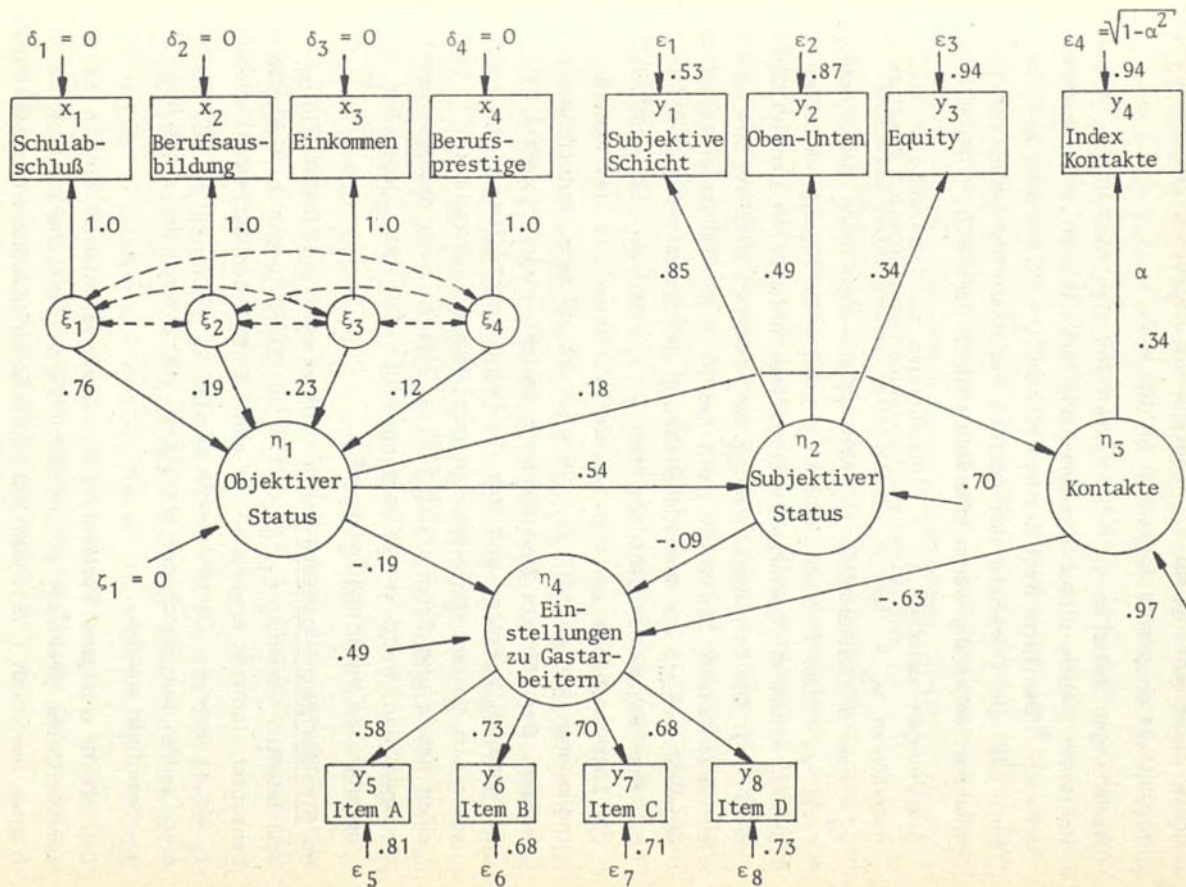
Damit führen wir über die Definition von Subgruppen die Dimension des Wettbewerbs bzw. der Wettbewerbserwartung als zentrale Determinante von Einstellungen zu Gastarbeitern in die Modelle mit ein.

8. Ergebnisse

LISREL-Modelle setzen sich aus zwei Teilen zusammen: Das Strukturgleichungsmodell als System theoretischer Sätze umfaßt die Beziehungen zwischen den latenten Variablen (η, ξ), das Meßmodell als System expliziter Korrespondenzhypthesen umfaßt die Beziehungen zwischen den latenten Variablen und ihren Indikatoren (x, y). Bei der sekundäranalytischen Auswertung sind dem Meßmodell durch die Notwendigkeit des Rückgriffs auf bereits vorgegebene Variablen enge Grenzen gesetzt.

Da wir die wahren exogenen Variablen ξ_1 bis ξ_4 mit den sie messenden Indikatoren x_1 bis x_4 gleichgesetzt haben, bleiben im Meßmodell folgende Beziehungen variabel:

- Beziehungen zwischen η_2 = "Subjektiver Status" und den Indikatoren y_1 = "Subjektive Schicht", y_2 = "Oben-Unten" und y_3 = "Equity": Der



stärkste Zusammenhang besteht zwischen der latenten endogenen Variablen η_2 und y_1 mit .85; für y_2 besteht mit .49 ein noch relativ deutlicher Zusammenhang mit η_2 , während die Beziehung zwischen η_2 und y_3 mit .34 als gering angesehen werden muß.⁷

- Beziehungen zwischen η_3 = "Kontakte" und dem Indikator y_4 = "Index Kontakte": Auch dieser Zusammenhang erweist sich mit .34 als sehr schwach.⁸ Der hohe Wert des Residuums (ζ_3 = .97) von η_3 weist darauf hin, daß zur Erklärung der latenten Variablen "Kontakte" weitere Indikatoren notwendig wären (Standardschätzfehler: .12).
- Beziehungen zwischen η_4 = "Einstellungen zu Gastarbeitern" und den Indikatoren y_5 = Item A (Gastarbeiter sollten ihren Lebensstil ...), y_6 = Item B (Gastarbeiter bei Arbeitsplatzmangel nach Hause schicken ...), y_7 = Item C (Gastarbeitern politische Betätigung verbieten ...) und y_8 = Item D (Gastarbeiter Ehepartner unter ihren Landsleuten suchen ...): Die Operationalisierung der latenten Variablen "Einstellungen zu Gastarbeitern" durch die Items des ALLBUS 1980 ist erheblich aussagekräftiger als die Operationalisierung der Kontakt-Variablen η_3 . Nur das eher allgemeiner gehaltene Item A (y_5) fällt mit .58 etwas zurück, die Zusammenhänge zwischen latenter Variablen und den anderen drei Indikatoren können mit .73, .70 und .68 als sehr deutlich bezeichnet werden. Der stärkste Zusammenhang besteht erwartungsgemäß zwischen der latenten Variablen und Item B ("Wenn Arbeitsplätze knapp werden, sollte man Gastarbeiter wieder in ihre Heimat zurückschicken"). Im Bereich Arbeit und Beruf treten unserer Ansicht nach Wettbewerbserfahrungen und Wettbewerbserwartungen der Befragten gegenüber Gastarbeitern am vordringlichsten auf.

Im Strukturgleichungsmodell haben wir einen unmittelbaren Einfluß der drei latenten Variablen η_1 , η_2 und η_3 auf Einstellungen zu Gastarbeitern behauptet. Indirekt erwarteten wir einen Einfluß der exogenen Variablen ξ_1 bis ξ_4 über η_1 . Daneben sollte η_1 über η_2 und η_3 auf η_4 wirken. Kontakte sollten nur durch den objektiven, nicht durch den subjektiven Status beeinflusst werden.

Die latente endogene Variable η_1 = "Objektiver Status" wird in der Gesamtstichprobe praktisch ausschließlich durch den Schulabschluß der Befragten beeinflusst (.76), während alle anderen Indikatoren von geringerer

Bedeutung sind. η_1 wiederum wirkt recht deutlich auf η_2 = "Subjektiver Status" (.54), aber nur schwach (.18) auf η_3 = "Kontakte".

Für die Gesamtstichprobe kann festgehalten werden, daß weder der objektive noch der subjektive Status der Befragten einen deutlichen Einfluß auf ihre Einstellungen zu Gastarbeitern haben. Die noch unzureichend operationalisierte Variable "Kontakte" wirkt sich mit -.63 am deutlichsten auf die Einstellungen zu Gastarbeitern aus. Wir vermuten, daß die Operationalisierung der Kontakte im ALLBUS 1980 zwar direkt nur auf die Tatsache faktischer Kontakte abzielt, daß aber die Einstellungen zu Gastarbeitern mit beeinflußt werden durch Erfahrungen, Erwartungen etc., die mit diesen tatsächlichen Kontakten zusammenhängen.

Inhaltlich heißt dies, daß ein deutlicher Zusammenhang besteht zwischen Kontakten zu Gastarbeitern und Einstellungen zu Gastarbeitern, und zwar derart, daß mit zunehmender Häufigkeit von Kontaktfeldern den Items des ALLBUS 1980 deutlich weniger zugestimmt wird. Je geringer aber die Zustimmung zu diesen Items ausfällt, um so geringer ist die Diskriminierung von Gastarbeitern. Das heißt, bei der Stichprobe des ALLBUS 1980 reduzieren Kontakte zu Gastarbeitern deren Diskriminierung (zumindest auf verbaler Ebene!).

Die Gültigkeit eines Modells für die zugrundeliegenden Daten wird von LISREL mittels des Likelihood-Ratio-Tests bestimmt und in Form des χ^2 -Wertes und des dazugehörigen Probability Level ausgedrückt. Abgesehen davon, daß der Likelihood-Ratio-Test als globaler Test für eventuell notwendige Modellmodifikationen wenig aufschlußreich ist, hat er darüber hinaus den Nachteil, daß er sehr sensitiv auf ein Ansteigen der Stichprobengröße reagiert. Da bei einer sehr großen Stichprobe jede Kausalstruktur von den Daten falsifiziert wird, kann der Likelihood-Ratio-Test leicht zu falschen Schlüssen führen, wenn das Modell für eine solche große Stichprobe überprüft werden soll.⁹

Als Maß für die Gültigkeit des Modells für die Stichprobe betrachten wir deshalb statt des χ^2 -Wertes und des Probability Level die Differenzen zwischen den erwarteten und den empirischen Korrelationen ($S - \Sigma$) (vgl. Schmidt 1977: 126ff.) Da in dieser Residual-Matrix nur eine einzige Korrelation $\geq .10$ auftritt (-.11), kann das Modell als für die Daten gut bestätigt angesehen werden (vgl. Schmidt 1979: 13).¹⁰

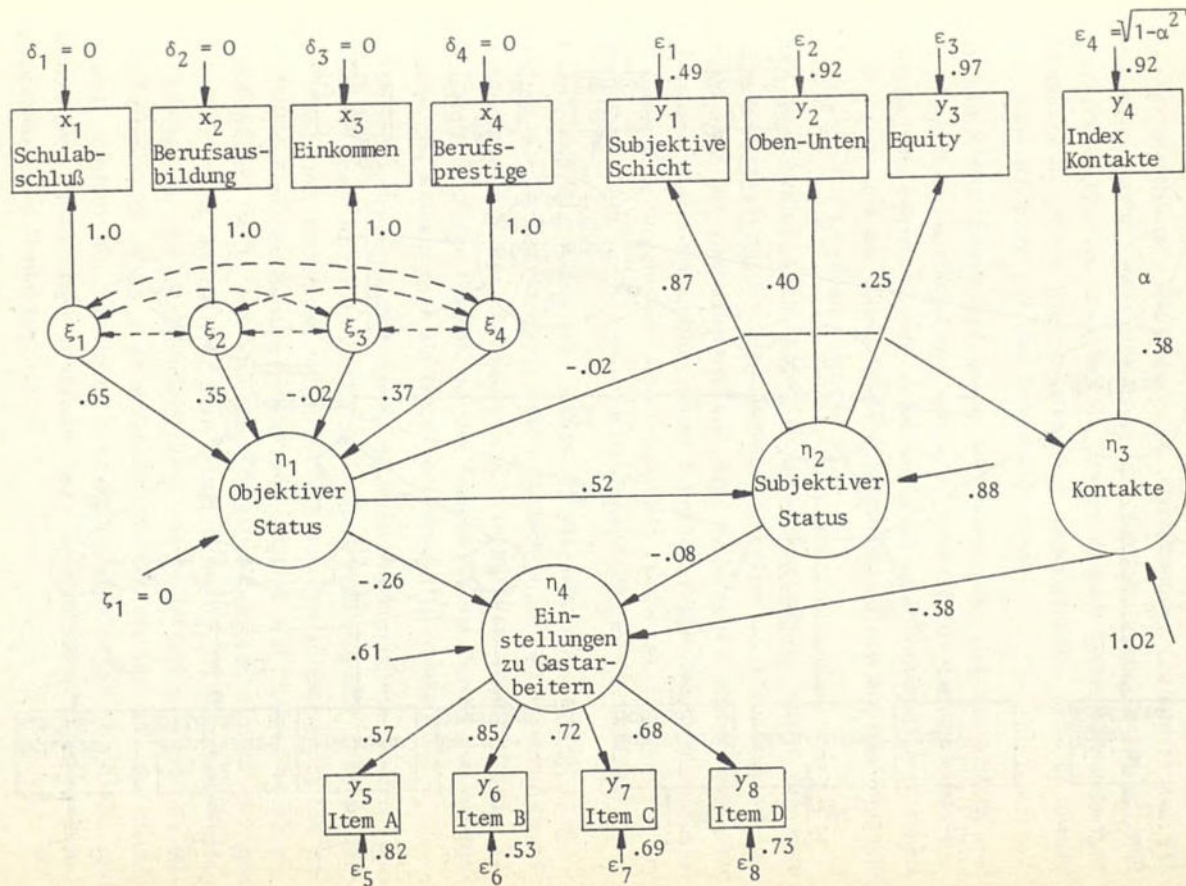
Dieses Modell überprüfen wir nun für verschiedene Subgruppen der Stichprobe, wobei wir über die Definition dieser Subgruppen die Dimension der Wettbewerbserfahrung bzw. Wettbewerbserwartung mit Gastarbeitern in die Analyse mit einbeziehen. Wir beschreiben dabei die Ergebnisse nicht mehr so detailliert wie bei Modell 1, sondern beschränken uns im wesentlichen auf die Darstellung auffälliger Unterschiede zwischen Subgruppen und Gesamtstichprobe und vor allem auf Gegensätze zwischen Subgruppen.

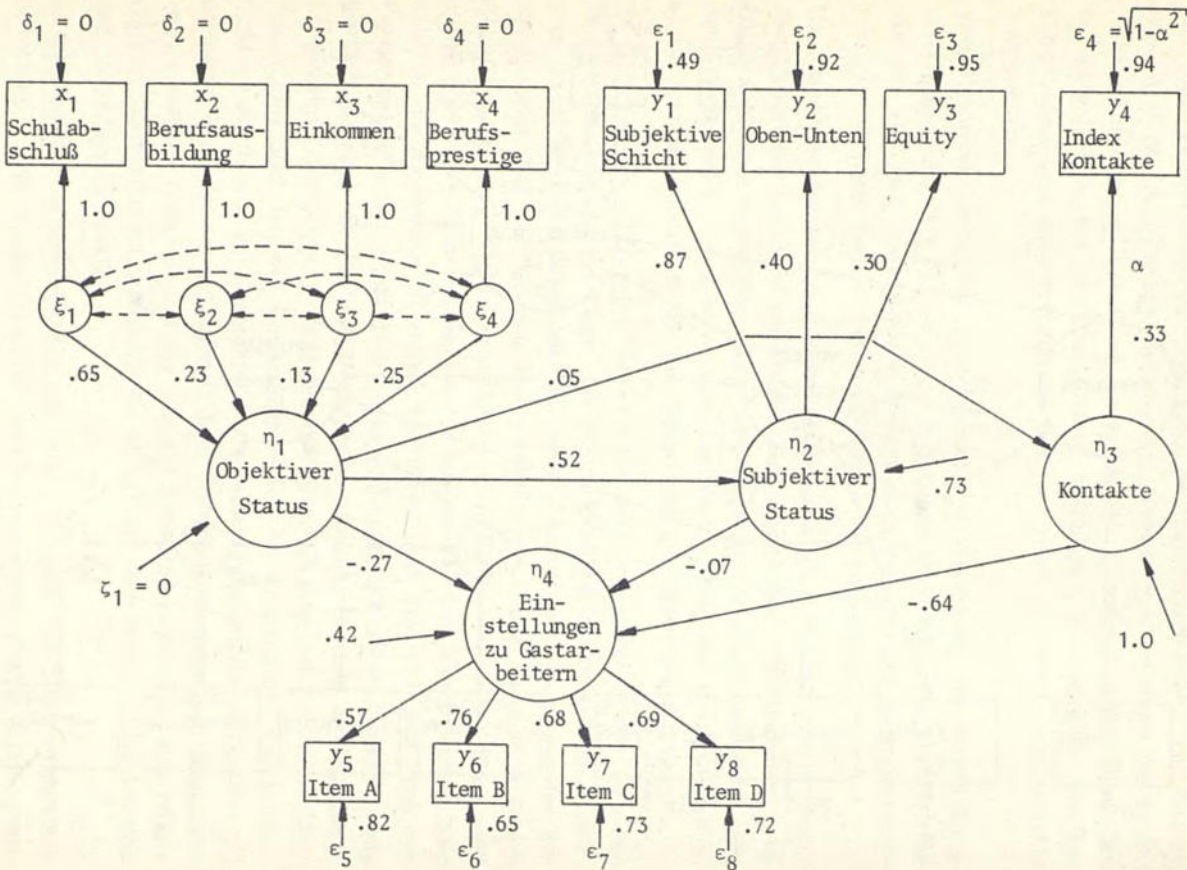
Vergleichen wir zunächst die Subgruppe mit Erfahrung von Arbeitslosigkeit (Modell 2a, $N = 293$) mit der Subgruppe ohne solche Erfahrung (Modell 2b, $N = 1908$), wobei über erstere die Dimension der Wettbewerbserfahrung mit Gastarbeitern in den Vergleich mit eingehen soll.

Bei Modell 2a bleiben in der Residual-Matrix ($S - \Sigma$) drei von insgesamt 75 Differenzen zwischen erwarteten und empirischen Korrelationen $\geq .10$. Damit kann das Modell für diese Subgruppe als bestätigt angesehen werden. Noch besser, ohne einen einzigen Wert $\geq .10$ in der Residual-Matrix, eignet sich das Modell 2b für die Subgruppe der "Nicht-Arbeitslosen".

Im Strukturgleichungsmodell fallen zwei deutliche Unterschiede zwischen den beiden Gruppen ins Auge. Bei der Bestimmung des objektiven Status, dies ist der erste Unterschied, spielen die Berufsausbildung und das Berufsprestige bei der Gruppe der "Arbeitslosen" eine deutlich stärkere Rolle als bei den "Nicht-Arbeitslosen", bei denen dafür das Einkommen etwas stärker auf den objektiven Status wirkt. Der zweite Unterschied betrifft die Beziehung zwischen Kontakten und Einstellungen: Bei den "Arbeitslosen" steht Kontakt in einem deutlich niedrigeren Zusammenhang mit Einstellungen als bei den "Nicht-Arbeitslosen". Ganz im Sinne unserer Erwartungen hinsichtlich Wettbewerbserfahrung wird bei "Nicht-Arbeitslosen" (wie auch bei der Gesamtstichprobe) mit zunehmender Kontakthäufigkeit (als Summe von Kontakten in unterschiedlichen Handlungsfeldern) geringere Diskriminierung verbalisiert, während "Arbeitslose" den Items relativ stärker (genauer: relativ weniger schwach) zustimmen, also relativ stärker diskriminieren. Dies könnte ein Effekt der Wettbewerbserfahrung der "Arbeitslosen" sein.

Im Meßmodell unterscheiden sich die beiden Gruppen vor allem hinsichtlich der Items B und C, der stärkste Unterschied ergibt sich, erwartungsge-





Modell 2b: "Nicht-Arbeitslose"

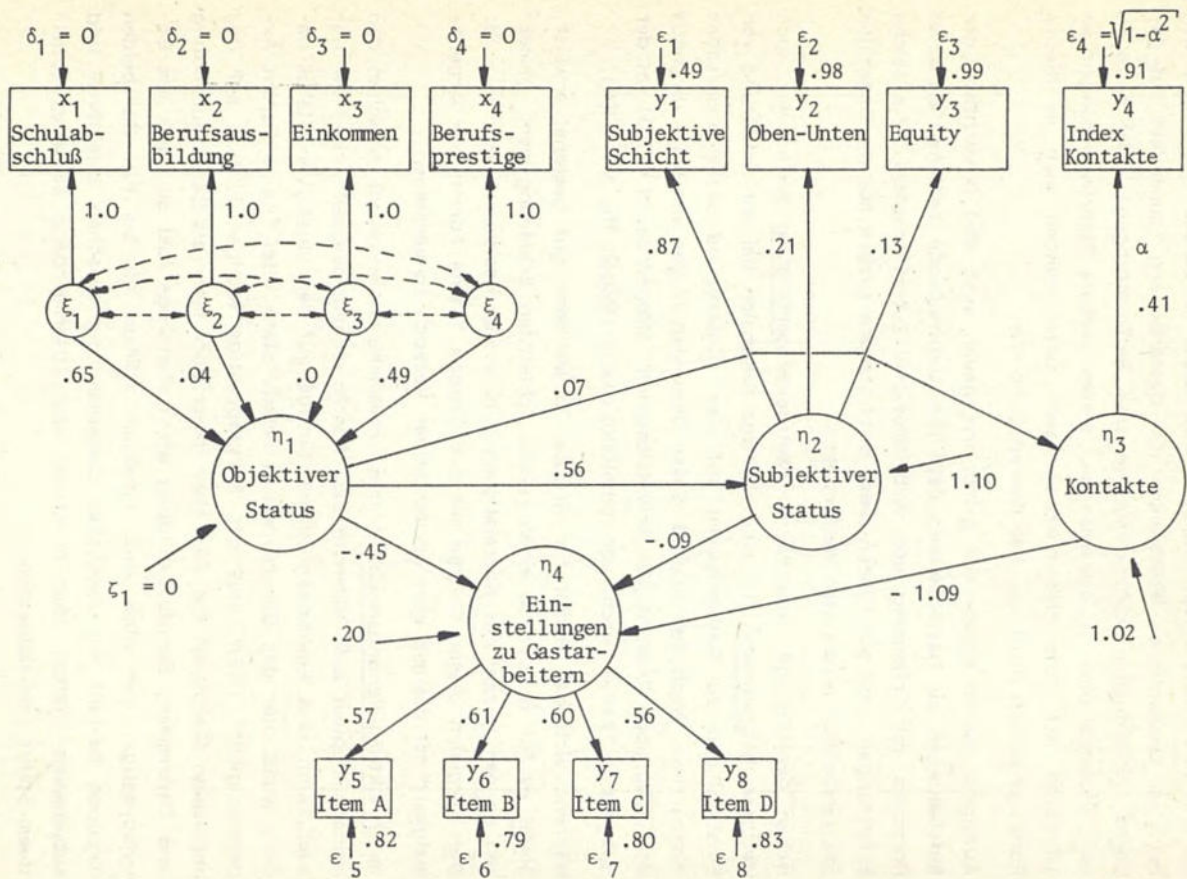
mäß, bei Item B. "Arbeitslose" reagieren offensichtlich zustimmender auf die Aussage, man solle Gastarbeiter wieder in ihre Heimat zurückschicken, wenn Arbeitsplätze knapp würden. Das heißt, im Bereich Arbeit und Beruf, wo tatsächlicher Wettbewerb mit Gastarbeitern (zumindest auf der Ebene individueller Wahrnehmungen und Befürchtungen) erfahren wird und allgemein eher zu erwarten ist, treten verbale Diskriminierungen am stärksten auf. Die "Nicht-Arbeitslosen" unterscheiden sich in diesem Punkt praktisch nicht von der Gesamtstichprobe.

Aufgrund dieser Ergebnisse gehen wir davon aus, daß hinsichtlich der Einstellungen zu Gastarbeitern deutliche Unterschiede bestehen zwischen Personen mit Erfahrung von Arbeitslosigkeit und Personen ohne solche Erfahrungen, und wir führen diese Unterschiede zurück auf die Dimension des tatsächlich erfahrenen Wettbewerbs.

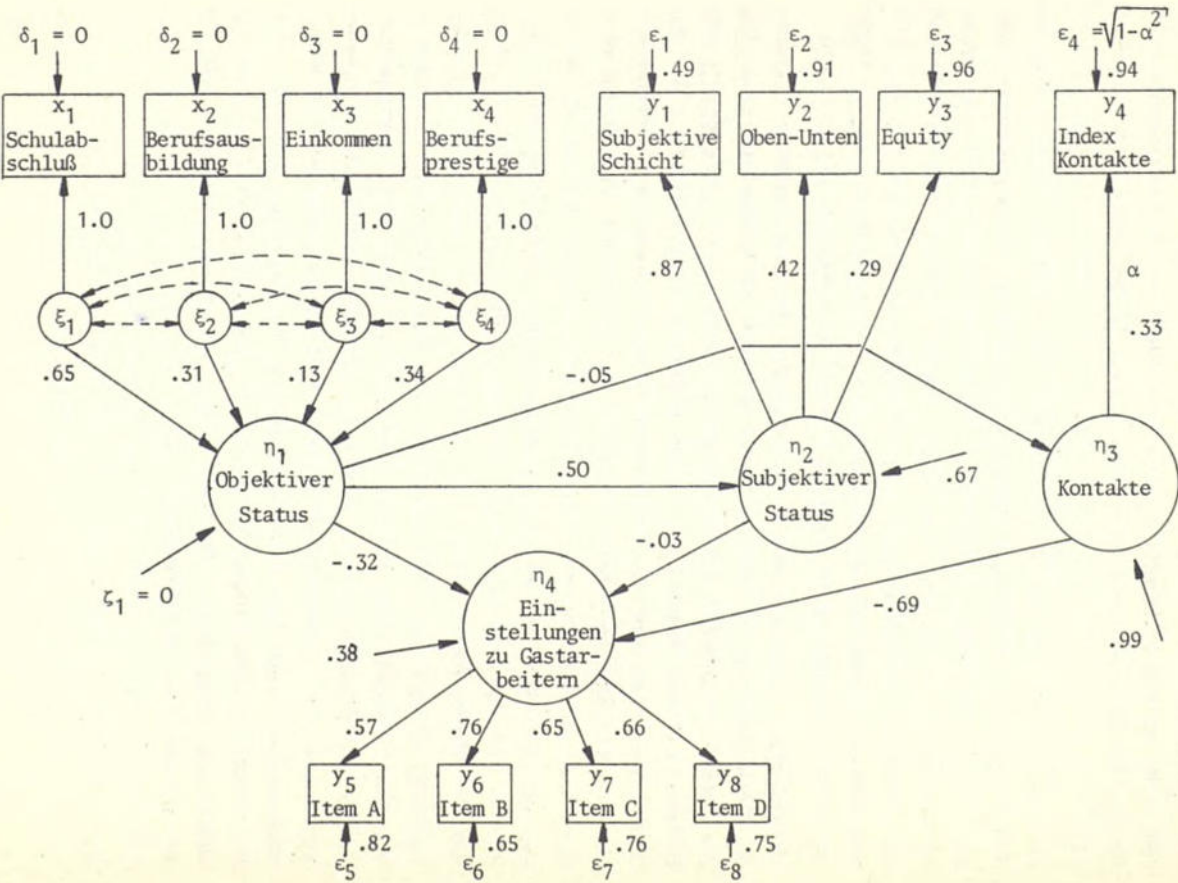
Neben dem Einfluß tatsächlicher Wettbewerbserfahrung halten wir auch Wettbewerbserwartung für eine wichtige Dimension bei der Erklärung von Einstellungen zu Gastarbeitern auf dem Hintergrund sozioökonomischer Konkurrenz. Nach dem Einfluß dieser Dimension fragen wir beim Vergleich der Subgruppe "Furcht vor Arbeitslosigkeit" (Modell 3a, N = 100) mit der Subgruppe "Keine Furcht vor Arbeitslosigkeit" (Modell 3b, N = 1309).

Während sich das Modell für Gruppe 3b als sehr gut geeignet erweist, bietet es für Gruppe 3a einen relativ schlechten Erklärungswert, gemessen an der Anzahl der Korrelationen $\geq .10$ in der Residual-Matrix (S - Σ). Der Vergleich dieser Gruppe mit der Gruppe "Keine Furcht vor Arbeitslosigkeit" ist deshalb eher als tentativer Versuch zu verstehen.

Im Strukturgleichungsmodell treten erhebliche Unterschiede zwischen den beiden Gruppen auf. Während bei "Furcht Arbeitslosigkeit" (FA) Berufsausbildung und Einkommen keinen Einfluß auf den objektiven Status haben, wirkt hier das Berufsprestige relativ stark. Bei "Keine Furcht Arbeitslosigkeit" (KFA) hat das Berufsprestige weniger Einfluß auf den objektiven Status als bei FA, aber immer noch mehr als Berufsausbildung und Einkommen; Berufsausbildung wirkt allerdings fast so stark wie Berufsprestige, vor allem aber erheblich stärker als bei FA. Bei beiden Gruppen besteht ein deutlicher Zusammenhang zwischen objektivem und subjektivem Status, aber praktisch kein Zusammenhang zwischen objektivem Status und Kontakten.



Modell 3a: "Fürcht vor Arbeitslosigkeit" (FA)



Modell 3b: "Keine Furcht vor Arbeitslosigkeit" (KFA)

Bei FA hat der objektive Status einen deutlich höheren Einfluß auf Einstellungen zu Gastarbeitern als bei KFA, während der subjektive Status praktisch ohne Einfluß bleibt, und zwar für beide Gruppen. Bei dem Wert von -1.09 für die Beziehung zwischen Kontakten und Einstellungen bei FA handelt es sich um eine "improper solution", die nicht interpretierbar ist.¹¹

Personen mit Furcht vor Arbeitslosigkeit, aber relativ hohem objektiven Status, diskriminieren Gastarbeiter möglicherweise deshalb weniger, weil sie in ihnen im Falle tatsächlicher Arbeitslosigkeit sicher keine echten Konkurrenten um Arbeitsplätze sehen. Dagegen ist der geringere Einfluß des objektiven Status auf die Einstellungen zu Gastarbeitern bei KFA möglicherweise dadurch zu erklären, daß sie eben gerade keine Furcht vor Arbeitslosigkeit haben, damit also auch keine Wettbewerbserwartung.

Im Meßmodell bestehen zwischen den beiden Gruppen sowohl hinsichtlich der Erklärung des subjektiven Status als auch hinsichtlich der Erklärung der latenten Variablen Einstellungen zu Gastarbeitern merkliche Unterschiede. Bei dem subjektiven Status macht sich die Einstufung auf der Oben-Unten-Skala und die Wahrnehmung gesellschaftlicher Gerechtigkeit bei der KFA-Gruppe erheblich stärker bemerkbar als bei FA. Die Einstellungen zu Gastarbeitern betonen bei KFA erheblich stärker als bei FA das Arbeitsplatz-Item und das Ehepartner-Item, dieses sogar leicht stärker als das Item der politischen Betätigung, das bei FA wichtiger ist als das Ehepartner-Item.

Während im Strukturgleichungsmodell das Zustandekommen der latenten Variablen Einstellungen zu Gastarbeitern im Sinne einer Wettbewerbs-Interpretation verstanden werden kann, sprechen die zuletzt genannten Ergebnisse des Meßmodells eher gegen eine solche Annahme. Wir weisen noch einmal darauf hin, daß die Ergebnisse mit Sicherheit durch die relativ schlechten Schätzungen des Modells für FA, bedingt durch die geringe Größe der Subgruppe, beeinflußt sind; dies ist zu berücksichtigen bei der Interpretation der Ergebnisse.

9. Zusammenfassung

Einstellungen zu Gastarbeitern werden im wesentlichen beeinflusst durch tatsächliche Kontakte zu ihnen. Daneben spielt der objektive, nicht aber der subjektive Status der Befragten eine Rolle bei der Erklärung solcher Einstellungen; zwischen unterschiedlichen Subgruppen treten allerdings von der Stärke der Beziehungen her gesehen deutliche Unterschiede auf.

Objektiver Status wurde empirisch bestimmt als latente Variable aus den gemessenen Variablen "Schulabschluß", "Beruflicher Ausbildungsabschluß", "Einkommen" und "Berufsprestige".

Der bewußte Verzicht auf die Konstruktion des objektiven Status als Index unter der Annahme gleicher Beiträge der einzelnen Variablen oder a priori behaupteter unterschiedlicher Beiträge hat sich bewährt. Zum einen hat sich gezeigt, daß die vier Indikatoren des objektiven Status sowohl bei der Gesamtstichprobe als auch bei allen Subgruppen gerade nicht mit gleichem Gewicht zur Bestimmung der latenten Variablen beigetragen haben. Zum andern konnte aufgezeigt werden, daß die einzelnen gemessenen Variablen selbst bei den unterschiedlichen Subgruppen in durchaus unterschiedlicher Weise zur Erklärung des objektiven Status beigetragen haben.

Auf der Ebene der latenten Variablen zeigte sich im großen und ganzen, daß Personen mit höherem objektiven Status und Kontakten zu Gastarbeitern in mehreren Kontaktbereichen den Einstellungssitems des ALLBUS 1980 deutlich schwächer zustimmen als Personen mit niedrigerem objektiven Status und weniger Kontakten, also zumindest auf verbaler Ebene eine geringere Diskriminierungsbereitschaft bzw. weniger verbales Diskriminierungsverhalten gegenüber Gastarbeitern zeigen.

Um der Frage nachzugehen, ob auch die Erfahrung bzw. die Erwartung sozioökonomischen Wettbewerbs, insbesondere in Form von Konkurrenz um Arbeitsplätze, Einfluß auf die Einstellungen zu Gastarbeitern haben, wurden LISREL-Gruppenvergleiche durchgeführt. Von der Annahme ausgehend, daß sowohl im Meßmodell als auch im Strukturmodell Unterschiede zwischen verschiedenen Subgruppen zu erwarten seien, hielten wir diese Vorgehensweise ebenso für angemessen wie für notwendig.

Subgruppenspezifische Analysen werden der Annahme gerecht, daß im Meßmodell, also in der Operationalisierung der Variablen, für unterschiedliche Subgruppen unterschiedliche Ergebnisse zu erwarten seien, bedingt etwa durch unterschiedliches Verständnis der Frageninhalte oder unterschiedliches Verbalisierungsverhalten der Angehörigen unterschiedlicher Subgruppen. Solche Unterschiede werden durch die Analyse der Ergebnisse für die Gesamtstichprobe allein vernachlässigt.

LISREL-Gruppenvergleiche haben darüber hinaus den bedeutenden Vorteil, daß man mit ihrer Hilfe nicht nur das Meßmodell, sondern auch das Strukturmodell, also die Beziehungen zwischen latenten Variablen, nach unterschiedlichen Ergebnissen für unterschiedliche Subgruppen analysieren kann.

Die empirischen Ergebnisse dieses Artikels sprechen denn auch dafür, über Analysen für Gesamtstichproben hinauszugehen und verstärkt subgruppenspezifische, vergleichende Analysen durchzuführen.¹²

Es zeigte sich nämlich, daß die Dimension der Wettbewerbserfahrung bzw. -erwartung im sozioökonomischen Bereich, d.h. hier vor allem die Konkurrenz um Arbeitsplätze, offensichtlich einen Einfluß hat auf Einstellungen zu Gastarbeitern. Zugegebenermaßen kann diese Frage hier nicht völlig eindeutig, sondern nur tendenziell beantwortet werden, da das zugrunde gelegte LISREL-Modell nicht bei allen Subgruppen einen optimalen Erklärungswert hatte. Soweit aber eine direkte Vergleichbarkeit zwischen Gruppen mit und ohne Wettbewerbserfahrung bzw. -erwartung gegeben war, deuten die Ergebnisse allerdings eindeutig in die Richtung eines Einflusses der Wettbewerbsdimension.

Negative Einstellungen und Diskriminierungen auf verbaler Ebene sind also offensichtlich kein repräsentatives Einstellungsmuster der bundesdeutschen Gesellschaft, sondern relativ häufiger bei Personen mit niedrigerem objektiven Status zu finden.

Erfahrungen mit Wettbewerb und Erwartungen von Wettbewerb im sozioökonomischen Bereich, vor allem als Konkurrenz um Arbeitsplätze, verstärken offensichtlich die Diskriminierungsbereitschaft, während tatsächliche Kontakte in jedem Falle eine wichtige Rolle bei der Verhinderung bzw. Reduzierung von Diskriminierungen spielen.

Anmerkungen

- 1 Wir danken Werner Hagstotz und Jürgen Hoffmeyer-Zlotnik (ZUMA) sowie Georgios Papastefanou (Sfb 3 der Universität Mannheim) für die Durchsicht des Manuskripts und eine Reihe kritischer und konstruktiver Hinweise. Unser besonderer Dank gilt Peter Schmidt (Universität Gießen), der mit einer intensiven LISREL-Beratung sowie tatkräftiger Unterstützung bei der Durchführung der Analysen erheblich am Zustandekommen dieses Artikels beteiligt war.
- 2 Geht man davon aus, daß ethnische Diskriminierungen in einem deutlichen Zusammenhang stehen mit der Wahrnehmung und Klassifizierung von Personen aufgrund äußerlicher und vor allem i.d.R. nicht veränderbarer Merkmale wie z.B. Hautfarbe, ist im übrigen nicht zu erwarten, daß die Anpassung an den Lebensstil der Gastberggesellschaft erheblichen Einfluß auf die Einstellungen ihrer Mitglieder zu Gastarbeitern haben dürfte.
- 3 So sprechen sich die SPD und der Deutsche Gewerkschaftsbund für, die CDU/CSU gegen ein kommunales Wahlrecht für Ausländer aus (vgl. DER SPIEGEL, 35.Jg. 1981, Heft 1/2: 28ff.).
- 4 Zur Bestimmung des Berufsprestiges s. Treiman, 1979; dort ist auch die "internationale Standardskala des Berufsprestiges" zu finden.
- 5 Zur Verknüpfung von objektiven und subjektiven Statusvariablen und deren Übersetzung in ein LISREL-Modell s. Sörböm und Jöreskog, 1981.
- 6 Aus fragebogentechnischen Gründen können bei dieser Kategorie nur solche Personen berücksichtigt werden, die zur Zeit ganztags oder halbtags hauptberuflich erwerbstätig sind, sowie Personen, deren Ehepartner zur Zeit halbtags oder ganztags hauptberuflich erwerbstätig ist.
- 7 Die Ursache für den geringen Erklärungswert von y_3 = Equity als Indikator für subjektiven Status muß in der Operationalisierung dieser Variablen gesehen werden. Die Wahrnehmung subjektiver Gerechtigkeit, also das Gefühl, weniger, gleichviel oder mehr "an den angenehmen Dingen des Lebens" zu erhalten als (in der Fragestellung nicht näher definierte) "Andere", leitet sich nicht zwangsläufig ab aus der subjektiven Positionswahrnehmung in einem vorgegebenen Schichtgefüge resp. einem vorgegebenen Oben-Unten-Kontinuum. Empirisch besteht im Gegenteil sogar eine relativ deutliche Unabhängigkeit zwischen subjektiver Schichteinstufung bzw. Oben-Unten-Einstufung und der Wahrnehmung subjektiver Gerechtigkeit (d.h. zum Beispiel, die Selbsteinstufung in die "Oberschicht" impliziert nicht per se die Wahrnehmung, im Vergleich zu anderen "mehr als den gerechten Anteil" zu erhalten). Wenngleich wir sowohl aus methodischen als auch empirischen Gründen vorschlagen, in weiteren Analysen auf die Verwendung der Variablen Equity als Indikator für subjektiven Status zu verzichten, behalten wir selbst diese Variable in den folgenden Modellen für Teilstichproben bei, da nicht von vornherein zu erwarten ist, daß sich den Ergebnissen für die Gesamtstichprobe entsprechende Ergebnisse auch bei Teilstichproben zeigen und ein wesentliches Anliegen des vorliegenden Artikels ja gerade die Frage nach der Notwendigkeit subgruppenspezifischer, vergleichender Analysen darstellt.

- 8 Der hohe Meßfehler ε_4 von y_4 weist auf eine hohe Fehlervarianz beim Index Kontakte hin. Die dementsprechend geringe Zuverlässigkeit des Index kann aber insofern vernachlässigt werden, als LISREL strukturelle Beziehungen unter Berücksichtigung der Korrektur der Höhe der Zufallsmeßfehler schätzt (vgl. dazu Werts und Linn 1972); Schlüsse auf der Ebene der latenten Variablen werden demzufolge durch den hohen Meßfehler nicht berührt.
- 9 Zur Kritik am Likelihood-Ratio-Test s. Schmidt (1977: 126ff.).
- 10 Im Modell angedeutet, aber nicht mehr näher ausgeführt, sind die Korrelationen der ξ -Variablen. Durch die Gleichsetzung von ξ_1 bis ξ_4 mit x_1 bis x_4 sind diese Korrelationen identisch mit den Pearson-Produkt-Moment-Korrelationen der gemessenen Variablen x_1 bis x_4 :

	$\xi_1(x_1)$	$\xi_2(x_2)$	$\xi_3(x_3)$	$\xi_4(x_4)$
$\xi_1(x_1)$	1.0			
$\xi_2(x_2)$.53	1.0		
$\xi_3(x_3)$.21	.37	1.0	
$\xi_4(x_4)$.18	.49	.36	1.0

- 11 Ebensowenig sinnvoll zu interpretieren sind die Residuen ξ_2 von η_2 mit einem Wert von 1.10 (Standardschätzfehler: .63) bzw. ξ_3 von η_3 mit einem Wert von 1.02 (Standardschätzfehler: .45). Methodisch gesehen handelt es sich bei solchen Werten um "improper solutions" bei der Maximum-Likelihood-Schätzung, bedingt durch die geringe Stichprobengröße; der kritische Wert liegt etwa bei $N = 200$ (vgl. dazu Boomsma, 1982).
- 12 Der nächste Schritt hätte nun in einem Konstantenvergleich zu bestehen, bei dem bestimmte Elemente des Modells über die Subgruppen hinweg stabil zu halten wären. Über die Ergebnisse in diese Richtung gehender Analysen wird zu einem späteren Zeitpunkt berichtet werden (in Vorbereitung: Krauth und Porst: Zur Erklärung diskriminierender Einstellungen zu Gastarbeitern, 1983).

Literatur

- Allport, Gordon W., 1954: The Nature of Prejudice. Cambridge, Mass.: Addison Wesley Publishing Company, Inc.
- Amir, Yehuda, 1969: Contact Hypothesis in Ethnic Relations. Psychological Bulletin, 71, Heft 5: 319-342.
- Boomsma, Anne, 1982: The Robustness of LISREL against Small Samples. In: Karl G. Jöreskog und H. Wold (Hrsg.), Systems under Indirect Observation. Amsterdam. (im Druck).
- Bundesminister für Arbeit und Sozialordnung (Hrsg.), 1981: Sozialpolitische Informationen XV/4, 3.März.
- Bundesminister für innerdeutsche Beziehungen (Hrsg.), 1974: Materialien zum Bericht zur Lage der Nation.

- Esser, Hartmut, 1980: Aspekte der Wanderungssoziologie. Darmstadt, Neuwied: Luchterhand.
- Hoffmann-Nowotny, Hans-Joachim, 1976: Gastarbeiterwanderungen und soziale Spannungen. S.43-62 in: Helga Reimann und Horst Reimann (Hrsg.), Gastarbeiter. München: Goldmann.
- Jöreskog, Karl G. und Dag Sörbom, 1978: LISREL IV. Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. User's Guide. Uppsala.
- Katsoulis, Harris, 1978: Bürger zweiter Klasse. Ausländer in der Bundesrepublik. Frankfurt am Main, New York: Campus.
- Krauth, Cornelia und Rolf Porst, 1983: Zur Erklärung diskriminierender Einstellungen zu Gastarbeitern (Arbeitstitel). (in Vorbereitung).
- Kühl, Jürgen, 1976: Die Bedeutung der ausländischen Arbeitnehmer für die Bundesrepublik Deutschland. S.23-42 in: Helga Reimann und Horst Reimann (Hrsg.), Gastarbeiter. München: Goldmann.
- Nikolinakos, Marios, 1974: Integration als Gleichberechtigung. Gewerkschaftliche Monatshefte, Heft 1: 20ff.
- Presse- und Informationsamt der Bundesregierung (Hrsg.), 1979: Gesellschaftliche Daten 1979.
- Reimann, Helga und Horst Reimann (Hrsg.), 1976: Gastarbeiter. München: Goldmann.
- Rose, A.M., 1973: Minderheiten. S.547-550 in: Wilhelm Bernsdorf (Hrsg.), Wörterbuch der Soziologie. Frankfurt am Main: Fischer-TB.
- Schäfer, Bernd und Bernd Six, 1978: Sozialpsychologie des Vorurteils. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Schmidt, Peter, 1977: Zur praktischen Anwendung von Theorien: Grundlagenprobleme und Anwendung auf die Hochschuldidaktik. Mannheim: Dissertation.
- Schmidt, Peter, 1979: On Decomposition of Effects in Causal Models. In: Materialien zum ZUMA-Workshop LISREL vom 8. bis 13. Oktober 1979. Mannheim.
- Sörbom, Dag und Karl G. Jöreskog, 1981: The Use of LISREL in Sociological Model Building. S.179-199 in: David J. Jackson und Edgar F. Borgatta (Hrsg.), Factor Analysis and Measurement in Sociological Research. A Multi-Dimensional Perspective. Beverly Hills.
- SPIEGEL, DER, 1975, 29. Jahrgang, Hefte 3, 18 und 50.
- SPIEGEL, DER, 1980, 34. Jahrgang, Heft 38.
- SPIEGEL, DER, 1981, 35. Jahrgang, Hefte 1 und 2.
- Statistisches Bundesamt Wiesbaden (Hrsg.), 1982: Wirtschaft und Statistik 1/82.
- Tomuschat, Christian, 1974: Die politischen Rechte der Gastarbeiter. S.80-101 in: Ansay Tugrul und Volkmar Gessner (Hrsg.), Gastarbeiter in Gesellschaft und Recht. München: Beck.
- Treimann, Donald J., 1979: Probleme der Begriffsbildung und Operationalisierung in der international vergleichenden Mobilitätsforschung.

S.124-167 in: Franz U. Pappi (Hrsg.), Sozialstrukturanalyse mit Umfragedaten. ZUMA-Monographien Sozialwissenschaftliche Methoden, Band 2. Königstein/Ts.: Athenäum.

Werts, C.E. und R.L. Linn, 1972: Corrections for Attenuation. Educational and Psychological Measurement, 33: 117-127.

SOZIALSTRUKTURELLE UND INDIVIDUELLE DETERMINANTEN VON SUBJEKTIVER SCHICHTIDENTIFIKATION UND POLITISCHEN EINSTELLUNGEN

Peter Schmidt und Gunter Wolf

1. Einführung

In diesem Aufsatz wird die empirische Umsetzung eines zentralen Problems der marxistischen und der soziologischen Theorie behandelt: Unter welchen sozialstrukturellen Bedingungen bildet sich bei Personen Klassenbewußtsein? Wie kann dieses theoretische Konzept expliziert und gemessen werden?

Nach einem kurzen Abriß der theoretischen Zusammenhänge wird das im ALLBUS 1980 verwendete Meßinstrument zur Erfassung der drei wichtigsten Kontaktpersonen der Befragten und deren persönlicher Merkmale dargestellt. Mit diesem Instrument können, über die Individualmerkmale der Befragten hinaus, sozialstrukturelle Merkmale ihrer primären Umwelt erfaßt und damit spezifische Kontexteffekte geprüft werden.

In dem sich anschließenden ersten empirischen Teil wird auf die Frage der Homogenität in den erfaßten Personennetzwerken, d.h. des Ausmaßes der Übereinstimmung in den Merkmalen der Befragten und deren Kontaktpersonen, eingegangen. Dabei wird die Konnektivität - das Ausmaß der wechselseitigen Beziehungen im Netzwerk - als wesentlicher Faktor berücksichtigt.

So wird zu prüfen sein, ob mit dem Ausmaß der Konnektivität der Personennetzwerke auch der Grad der Homogenität im Bezug auf spezifische Merkmale zunimmt.

Im nächsten Abschnitt wird ein komplexes Strukturgleichungsmodell spezifiziert, welches den Zusammenhang zwischen sozialstrukturellen Merkmalen des Befragten und seiner wichtigsten Kontaktpersonen (Kontextmerkmal) einerseits und subjektiver Schichtidentifikation sowie politischer Einstellungen des Befragten andererseits enthält. Mit Hilfe des Modells wird quantitativ der Effekt jeweils einer Einflußgröße - wie z.B. der beruflichen Stellung des Befragten - auf eine andere - wie Wahrnehmung

von Klassenkonflikten - unter Konstanthaltung der anderen Variablen des Modells bestimmt.

Den Abschluß bildet die Überprüfung dieses Modells für vier Subgruppen, die sich danach unterscheiden, wie viele der Kontaktpersonen auch jeweils untereinander befreundet sind. Durch diese Vorgehensweise soll geprüft werden, ob die Effekte in unterschiedlich stark verbundenen Netzwerken gleich sind, d.h. ob Kontexteffekte additiv wirken.

2. Theorie: Falsches Bewußtsein, subjektive Realität und Kontexteffekte: Abschied vom Proletariat?

Ein zentraler Gegenstand der soziologischen Theoriebildung und der Sozialstrukturanalyse ist bis heute der Zusammenhang zwischen sozialstrukturellen Merkmalen einer Person und den Merkmalen ihres sozialen Kontextes einerseits, und andererseits von subjektiver Schichtidentifikation, Wahrnehmung von Klassenkonflikten und politischen Einstellungen der Person selbst (vgl. Gorz 1971). Wenn Diskrepanzen zwischen diesen objektiven und subjektiven Merkmalen auftreten, spricht man von 'falschem Bewußtsein'. Die Pionierarbeit zu diesem Thema war zweifellos der "18. Brumaire ..." von Karl Marx, in dem er die politische und soziale Situation der Bauern im Frankreich des 19. Jahrhunderts untersuchte. Marx argumentiert, daß nur die Existenz eines Klassenbewußtseins eine Klasse für sich schafft. Wenn hingegen eine Gruppe mit gleichem ökonomischen und sozialen Hintergrund kein Klassenbewußtsein entwickelt, so besteht nur eine latente Klasse, eine Klasse an sich.

Ein wesentlicher Faktor ist nach Marx sowohl das individuelle Merkmal der Klassenzugehörigkeit wie auch die Art des sozialen Kontextes:

"Die Parzellenbauern bilden eine ungeheure Masse, deren Glieder in gleicher Situation leben, aber ohne in mannigfache Beziehung zueinander zu treten. Ihre Produktionsweise isoliert sie voneinander, statt sie in wechselseitigen Verkehr zu bringen. Die Isolierung wird gefördert durch die schlechten französischen Kommunikationsmittel und die Armut der Bauern.

Die Parzelle, der Bauer und die Familie; daneben eine andre Parzelle, ein anderer Bauer und eine andre Familie. Ein Schock davon macht ein Dorf, und ein Schock von Dörfern macht ein Department. So wird die große Masse der französischen Nation gebildet durch einfache Addition gleichnamiger Größen, wie etwa ein Sack von Kartoffeln einen Kartoffel-

sack bildet. Insofern Millionen von Familien unter ökonomischen Existenzbedingungen leben, die ihre Lebensweise, ihre Interessen und ihre Bildung von denen der anderen Klassen trennen und ihnen feindlich gegenüberstellen, bilden sie eine Klasse. Insofern ein nur lokaler Zusammenhang unter den Parzellenbauern besteht, die Dieseligkeit ihrer Interessen keine Gemeinsamkeit, keine nationale Verbindung und keine politische Organisation unter ihnen erzeugt, bilden sie keine Klasse." (Marx 1976: 117-118).

Neben der individuellen Zugehörigkeit zu einer Klasse ist damit der soziale Kontext, z.B. in Form von Isolation, von großer, wenn nicht entscheidender Bedeutung für die Ausprägung eines Klassenbewußtseins.

So begünstigt, im Gegensatz zu den Parzellenbauern, deren Produktionsweise sie untereinander isoliert, die Lebenssituation der Proletarier deren Kommunikation. Dieser Effekt tritt aus drei Gründen auf: Zunächst, weil sie in immer größeren Betrieben zusammenarbeiten und in den gleichen Regionen der Städte leben, sodann weil der Kapitalismus Kommunikationsmittel in bisher unbekannter Weise entwickelt hat, die auch den Proletariern zugänglich sind, und schließlich weil Organisationen der Arbeiter für Kommunikation sorgen oder diese zumindest erleichtern. Auf diese Weise können die Mitglieder des Proletariats wahrnehmen, daß ihre Leiden die Leiden aller Proletarier sind, daß das ihre Lebenslage verschlechternde Verhalten einzelner Kapitalisten das Verhalten aller Kapitalisten, und daß der Konflikt mit einzelnen Unternehmern ein Konflikt mit der die staatlichen Machtmittel kontrollierenden Klasse ist (vgl. Lindner 1972: 39).

Der Begriff des Klassenbewußtseins ist bei Marx, aber auch bei späteren Autoren wie Lukacs, nicht klar abgegrenzt. Giddens (1979) hat hierfür eine Explikation vorgeschlagen, die uns zur Klärung des Begriffs und zur Spezifikation unseres Modells nützlich erscheint.

Giddens (1979: 137) unterscheidet drei Stufen des Klassenbewußtseins:

1. Die am wenigsten entwickelte Form ist die der Klassenidentität und -differenzierung.
2. Die darauf folgende Stufe ist durch die Wahrnehmung von Klassenkonflikten charakterisiert.
3. Die dritte Stufe ist das revolutionäre Klassenbewußtsein, welches dadurch gekennzeichnet ist, daß eine grundsätzliche Umwälzung der Machtstruktur der Gesellschaft für möglich gehalten wird und die Ein-

stellung, daß Aktionen einer Klasse dies erreichen können. Da wir hierfür keine Daten im ALLBUS haben, gehen wir nicht näher auf diese dritte Stufe ein.

Im Gegensatz zu Marx argumentiert Giddens, daß auch subjektive Schichtidentifikation nicht deterministisch zur Wahrnehmung von Klassenkonflikten und diese wiederum nicht automatisch zu revolutionärem Klassenbewußtsein führt (vgl. hierzu Giddens 1979).

Die bisher formulierten Hypothesen sollen zusammenfassend reformuliert werden (vgl. Jackman und Jackman 1973, Kort-Krieger 1982, Schmidt 1982). Hierbei verwenden wir nicht Klasse, sondern objektiven Status als unabhängige Variable. Objektiver Status wird hierbei nicht von der Dichotomie Kapitalbesitzer-Arbeitnehmer wie bei Marx, sondern durch berufliche Stellung, Einkommen und Bildung bestimmt (vgl. hierzu Giddens 1979):

- H₁ Je höher der objektive Status einer Person, desto höher ihre subjektive Schichtzuordnung.
- H₂ Je höher die berufliche Stellung der Personen der primären Umwelt, repräsentiert durch die drei besten Freunde einer Person, desto höher ist ihre subjektive Schichtzuordnung.
- H₃ Je höher die subjektive Schichtidentifikation, desto geringer die Wahrnehmung von Klassenkonflikten.

Behandeln wir nun näher die bisherigen Studien und Hypothesen über den Zusammenhang zwischen subjektiver Schichtidentifikation und politischen Einstellungen.

Die zentrale Fragestellung (Centers 1949, Converse 1964) war folgendermaßen: Führt ein hohes Ausmaß an subjektiver Klassenidentifikation bzw. Klassenbewußtsein auch zu extremen politischen Ansichten?

Hierbei wurden folgende Hypothesen formuliert:

- H₄ Personen, die sich zur Arbeiterklasse zählen, tendieren dazu, sich politisch links einzustufen, während Personen, die sich stark mit der Mittelschicht identifizieren, sich eher politisch rechts einstufen.
- H₅ Personen, die sich zur Arbeiterklasse zählen, tendieren zu einer kollektiven Philosophie, während Personen, die sich zur Mittelschicht zählen, eher zu einer individualistischen Philosophie tendieren.

Centers operationalisierte das Konstrukt "kollektive Philosophie" mit Fragen nach Unterstützung für öffentliche Hilfen und Programme für sozial schwache Gruppen, Sympathie für Gewerkschaften u.a.m. Als weitere Hypothesen lassen sich formulieren:

H_6 Je höher die objektive Schicht einer Person, desto weniger deprimiert fühlt sie sich.

H_7 Je höher die subjektive Schicht einer Person, desto weniger deprimiert fühlt sie sich.

Kort-Krieger (1982) hat für H_1 und H_2 sowie zusätzliche Hypothesen ein komplexes LISREL-Modell spezifiziert und getestet. Schmidt (1982) hat dies im Anschluß daran für den Zusammenhang zwischen subjektiver Schicht und politischen Einstellungen ($H_3 - H_7$) durchgeführt. In diesem Beitrag sollen nun die beiden Modelle integriert und simultan getestet werden. Damit soll gleichzeitig der Zusammenhang zwischen objektiver und subjektiver Schicht sowie objektiver Schicht und politischen Einstellungen untersucht werden.

Darüber hinaus soll aber ein Aspekt besonders untersucht werden, den Marx im 18. Brumaire bereits stark betont hat, der aber bisher in der empirischen Sozialforschung durchweg zu kurz kam: der Einfluß des Kontextes. Durch die Erfassung der drei besten Freunde, einige ihrer sozialstrukturellen Merkmale sowie von Merkmalen des Netzwerkes der Freunde selbst können zumindest einige Aspekte des Einflusses des sozialen Kontextes mit Daten des ALLBUS näher analysiert werden. Gegenüber den meist verwendeten aggregierten Mittelwerten als Kontextvariablen hat die hier gewählte Vorgehensweise den Vorteil, relevante Aspekte des Kontextes der jeweils befragten Person zu erfassen (vgl. Esser 1982: 163-165).

3. Die Erfassung der Freunde

Die Berücksichtigung der primären Umwelt der Befragten und die Untersuchung der Einflüsse dieser Faktoren des sozialen Kontextes auf die Befragungsperson läßt sich bis in die frühen Wahlstudien zurückverfolgen (Katz und Lazarsfeld 1955, Feldarbeit 1945) und hat mit dem Aufschwung der Netzwerkanalyse wieder verstärkt an Bedeutung gewonnen.

Das verstärkte Wiederaufgreifen dieses Ansatzes hat sich jedoch nur in geringen Veränderungen des bereits anfangs benutzten Fragenformates niedergeschlagen.

Die Gewinnung von Daten des sozialen Kontextes in Repräsentativbefragungen, so scheint es, erfordert eine spezifische, inzwischen quasi standardisierte, Vorgehensweise der Erfassung ego-zentrierter Netzwerke (vgl. Laumann 1973, Pappi 1973, Laumann und Pappi 1976).

Auch in unserer Analyse sollen nicht nur die persönlichen Merkmale der Befragten untersucht werden, sondern der soziale Kontext soll gleichfalls - als Faktor, dem besondere Bedeutung beizumessen ist - einbezogen werden. Dieser Kontext wird repräsentiert durch die im Interview erfragten Freundschaftsbeziehungen der Befragten und die Angaben, die der Befragte über diese Freunde gemacht hat.

Oft wird das geringe Ausmaß von Kontexteffekten festgestellt, andererseits aber mit Recht darauf hingewiesen, daß damit noch nicht belegt ist, daß Kontexteffekte keine Wirkungen auf individuelle Akteure haben (vgl. z.B. Esser 1982: 163). Ein Grund dürfte darin liegen, daß die Kontextvariablen im allgemeinen pro Organisation, Gruppe oder administrative Einheit als Mittelwerte gebildet werden. Dies bedeutet aber, daß nicht unbedingt die individuell bedeutsamen Kontexte und Interaktionsnetzwerke erfaßt werden, die für die einzelne Person handlungsrelevant sind. Die im ALLBUS verwendete und nachstehend kurz beschriebene Erfassung der primären Umwelt ist eine Lösungsmöglichkeit dieses Problems.

Die generellen methodischen und theoretischen Probleme der Kontext- und Mehrebenenanalyse werden hier nicht näher behandelt. Man vergleiche hierzu Boyd und Iversen (1979), Clar (1981), Hummell (1974), Erbring und Young (1980).

Die im ALLBUS 1980 (vgl. Anhang 1) gewählte Fragestellung erfaßt den sozialen Kontext des Befragten über die von ihm erfragten drei besten Freunde. Dies ist eine Vorgehensweise, die sich mittlerweile durchgesetzt hat, aber durch die gewählte Beschränkung auf drei Freunde auch eine Eingrenzung der erfaßbaren primären Umwelt darstellt.

Ein Befragter kann weniger als drei Freunde haben, jedoch auch deutlich mehr. Während im ersten Fall einfach keine Angaben - zu nicht vorhan-

denen Freunden - erfolgen, wird im letzteren Fall der Befragte im Interview quasi zu einer Auswahl gedrängt. Hier kann für die Analyse nur angenommen werden, daß der Befragte die für ihn wichtigsten Freunde auch tatsächlich angegeben hat.

Nach den Ergebnissen des ALLBUS 1980 hat der weitaus überwiegende Teil der Befragten einen Freundeskreis, der zumindest drei Freunde umfaßt. Von den 2 955 Befragten hatten 340 überhaupt keinen Freund angegeben. Weitere 306 Befragte hatten lediglich einen Freund A angegeben, und 508 Befragte haben nur zwei Freunde genannt (Freund A und B).

Somit haben 1 801, d.h. 60,92% der Befragten, drei Freunde (Freunde A, B, C) angegeben. Zu den genannten Freunden wurden persönliche Merkmale erfragt und es wurde auch erfaßt, ob sich diese Freunde untereinander gut kennen bzw. untereinander befreundet sind. Diese Angabe des Befragten gibt uns zumindest einen Anhaltspunkt über die Beziehung selbst und ihre interne Struktur oder Dichte.

Die Gruppe der Befragten, die drei Freunde genannt hatten, wird von uns im weiteren für unsere Analyse benutzt werden. Sie scheint uns am geeignetsten, um einen zentralen Aspekt der Freundesnetze, deren interne Struktur oder Dichte, näher zu untersuchen.

Da wir im weiteren die Angaben des Befragten zu seinem erfaßten sozialen Kontext analysieren werden, ist die Frage nach der Validität und Zuverlässigkeit der Angaben wichtig. Wie in jeder Befragung oder Test-Retest-Situation ist zunächst ein gewisses Maß an Abweichung zwischen den erfaßten Daten und der Realität zu erwarten. Dies gilt für die Angaben des Befragten über sich selbst, wie auch für dessen Angaben über seine Freunde. Die Frage nach z.B. dem Einkommen oder aber der Wahlentscheidung wird nicht immer voll wahrheitsgemäß beantwortet werden. Trotz der generellen Freiwilligkeit der Teilnahme an der Befragung und der Möglichkeit der Verweigerung der Antwort sind gemachte Angaben nicht immer als voll der Realität entsprechend zu werten.

Diese fehlerbehafteten Angaben betreffen - im Normalfall einer Befragung - jedoch immer nur den Befragten selbst.

Bei der Erfassung ego-zentrierter Beziehungen bzw. primärer Umwelten gibt der Befragte jedoch auch Auskunft über andere Personen, wobei

diese Angaben Gegenstand der Analyse sind und meist als Realität hingenommen werden müssen. "Follow-Up"-Interviews mit den genannten Personen der primären Umwelt des Befragten sind im Normalfall nicht möglich, und die genannten Personen sind in der Regel auch nicht selbst Befragte, was eine Überprüfung der Angaben zuließe.

So bleibt im Normalfall die Situation, daß die Angaben der Befragten über sich selbst und die genannten Freunde als voll zutreffend betrachtet werden müssen, auch wenn empirische Ergebnisse Zweifel an dieser notwendigen Grundannahme aufwerfen (vgl. z.B. Bernard et al. 1981). Diese Meßfehler können sowohl zufällig - somit vergleichsweise weniger bedeutsam - als auch systematisch sein. Systematische Meßfehler betreffen dabei vor allem die Wahrnehmung von Merkmalen und Einstellungen geringer Sichtbarkeit (vgl. Scheuch 1965), während andererseits insbesondere demographische Merkmale weitgehend zutreffend erfaßt werden, und daher vom Befragten auch weitgehend zuverlässige Angaben über diese Merkmale seiner Freunde möglich und zu erwarten sind. Dies ist eines der Ergebnisse einer Analyse der Fehlerbehaftetheit von Angaben von Befragten über deren Freunde, wobei die Angaben der Freunde über sich selbst zur Kontrolle herangezogen werden konnten (Pappi und Wolf 1981).

Theoretisch bedeutet dies, daß die erklärenden Variablen hier nicht die Merkmale der Freunde selbst sind, sondern die Wahrnehmung des Befragten von den Merkmalen seiner drei besten Freunde. Da für die Erklärung des Befragtenverhaltens dessen Wahrnehmung verhaltensrelevant ist, ist es angemessen, die hier gemessenen Größen zu verwenden. Allerdings muß dabei beachtet werden, daß man nicht deterministisch auf die Merkmale der Freunde zurückschließen kann. Dies sind konzeptuell andere Größen, die durch explizite Anschlußtheorien mit der Wahrnehmung des Befragten verknüpft werden müssen. Die anderen in den Hypothesen vorher angesprochenen Variablen wie subjektive Schicht, Wahrnehmung von Klassenkonflikt etc. finden sich mit genauer Operationalisierung im Codebuch des ALLBUS 1980 und sollen hier nicht näher besprochen werden.

4. Empirische Analysen zur Struktur der Freundschaftsbeziehungen

Die Struktur der im Interview genannten Freundschaftsbeziehungen kann ausgehend von zwei Extremsituationen betrachtet werden, der segmentierten Beziehung, die nur durch und über den Befragten als zentrale Person hergestellt wird, und andererseits der voll verbundenen Beziehung, in der nicht nur der Befragte mit allen Freunden verbunden ist, sondern sich auch alle Freunde direkt untereinander gut kennen bzw. befreundet sind.

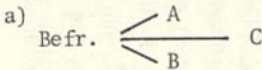
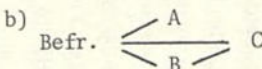
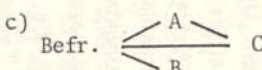
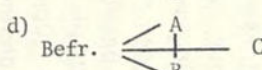
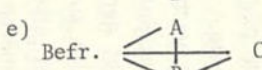
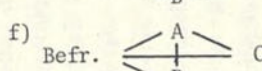
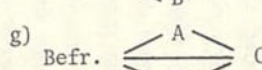
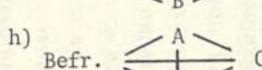
Eine Übersicht über die vorgefundene Verteilung dieser Beziehungstypen zeigt Tabelle 1. Die ermittelten Freundesnetze sind nur vergleichsweise selten (12% der Fälle) in dem Sinne segmentiert, daß sich die Freunde untereinander überhaupt nicht kennen. Die beiden Mischtypen, überwiegend unverbundene Netze - nur zwei Freunde kennen sich untereinander - bzw. überwiegend verbundene Netze - zwischen den drei Freunden bestehen zwei direkte Beziehungen - treten bereits deutlich häufiger auf. Deutlich am stärksten ist der Beziehungstypus des voll verbundenen Netzes vertreten, in dem sich auch alle genannten Freunde untereinander gut kennen oder befreundet sind. Über die Hälfte aller hier untersuchten Freundesbeziehungen gehören diesem Typus an.

Die Interpretation dieser Verteilung der Beziehungstypen wird durch die Art der Fragestellung im Interview erschwert. Erfragt wurden die drei besten Freunde des Befragten, jedoch haben wir keinen Indikator für die Stärke der Beziehung oder die subjektive Bedeutung, die der Befragte dieser - durch die Befragung geforderten - Rangreihe zugrunde legt. So können wir die angegebenen Beziehungen nur als gleich gewichtig bzw. für den Befragten gleich bedeutsam - und weiterhin nur als symmetrisch - annehmen.

Wir wollen nun versuchen, einen Einblick in die erfaßten Beziehungen selbst zu gewinnen, indem wir einige der erfragten Merkmale der Freunde mit den Merkmalen der Befragten vergleichen.

Da das Ziel unserer Untersuchung die Betrachtung der Struktur der Beziehung ist, die die vier Personen zusammen bilden, müssen wir von der Betrachtungsebene der Beziehung insgesamt ausgehen.

Tabelle 1: Übersicht über die Art und Verteilung der Freundesbeziehungen

Typus der Freundes- beziehung	Beschreibung des Be- ziehungstypus	Graphische Darstellung	Index der Dich- te der Beziehung	Anzahl der Fälle
1. voll segmentierte Netze	Kontakt läuft aus- schließlich über den Befragten ab.	a) 	3	213
2. Überwiegend segmen- tierte Netze	Kontakt läuft über- wiegend über und durch den Befragten ab. Ein Freundes- paar kennt sich untereinander	b) 	4	453
		c) 		
		d) 		
		e) 		
3. Überwiegend dicht verbundene Netze	Kontakt läuft nicht nur über den Befrag- ten ab. Zwei Freun- despaare kennen sich auch untereinander.	f) 	5	158
		g) 		
		h) 		
4. Voll verbundene Netze	Kontakt verläuft zwischen dem Befrag- ten und allen Freunden direkt ab.		6	954
				1778

Es zeigt sich, daß der größte Teil der genannten Freundesnetze entweder voll gleichgeschlechtlich strukturiert ist (30,2%) - alle vier Personen haben dann das gleiche Geschlecht - oder aber zumindest überwiegend gleichgeschlechtlich ist (47,6%) - wobei drei der vier Personen der Beziehung dem gleichen Geschlecht angehören. Vollständig gemischte Beziehungen, d.h. von den vier Personen des Freundesnetzes sind je zwei männlich, je zwei weiblich, treten nur zu 22,2% auf. Betrachten wir diese Ergebnisse in Relation zu der Dichte der Beziehung, so findet sich nur ein hervorhebenswerter Effekt: voll segmentierte Beziehungen sind am stärksten von allen Beziehungstypen voll gleichgeschlechtlich (39,4%) oder aber überwiegend gleichgeschlechtlich (43,7%) strukturiert.

Differenziert man nach dem Geschlecht des Befragten, so zeigt sich, daß männliche Befragte zu 76,4% ein voll oder aber zumindest überwiegend männliches Freundesnetz haben. Neben dem Befragten sind alle, oder zumindest zwei der drei genannten Freunde männlich. Der gleiche Effekt ist festzustellen, sofern der Befragte weiblich ist: Die genannten Freunde sind ausschließlich, zumindest aber überwiegend, gleichfalls weiblichen Geschlechts. Ausgehend von der Dichte der Freundesbeziehung zeigt sich auch bei dieser Betrachtungsebene der bereits angeführte Effekt: voll segmentierte Beziehungen sind deutlich am stärksten voll gleichgeschlechtlich - und zwar unabhängig vom Geschlecht des Befragten - strukturiert.

Als zweites Merkmal von Befragten und Freunden soll auf deren politische Präferenzen am Beispiel der Wahlentscheidung eingegangen werden.

Für die Befragten wurden diese mittels der Sonntagsfrage ermittelt, für die genannten Freunde können wir uns nur auf die Angaben der Befragten über diese Freunde und deren politische Präferenzen stützen. Die Befragten wurden gefragt, welche Partei deren Freunde gewöhnlich wählen.

Diese Parteipräferenzen betrachten wir nur dichotomisiert in CDU/CSU und die - zum Befragungszeitraum bestehende - Regierungskoalition aus SPD und FDP. Präferenzen für andere Parteien - wobei allenfalls auf die "Grünen" nennenswerte Nennungen entfielen - können vernachlässigt werden.

Zunächst zum generellen Bild, das sich abzeichnet. Die Hälfte der Freundesnetze (51,2%) ist im Bezug auf politische Präferenzen vollständig ho-

mogen: Alle vier Personen teilen die Präferenz entweder für CDU/CSU oder SPD/FDP. Weitere 29,9% sind überwiegend homogen, d.h. drei der vier Freunde teilen die gleiche Parteipräferenz, nur einer der Freunde hat eine abweichende Präferenz.

Vollständig gemischte Präferenzen treten demgegenüber nur in 13,9% der Freundesnetze auf.

Auf einen Sonderfall soll noch kurz eingegangen werden. Theoretisch ist es möglich, daß der Befragte mit seiner Parteipräferenz innerhalb der gesamten Freundesbeziehung allein steht, seine Freunde also alle die jeweils andere politische Einstellung teilen. Nach unseren Daten handelt es sich tatsächlich um einen nur selten auftretenden Fall. Nur in 5,0% der Freundesnetze steht der Befragte nach seinen Angaben mit seiner Parteipräferenz allein und einem geschlossen politisch anders orientierten Freundeskreis gegenüber.

Stehen nun diese Ergebnisse in einem Zusammenhang mit der Struktur der Freundesbeziehung? Mit der Dichte der Beziehung, so können wir feststellen, steigt auch die vorgefundene Übereinstimmung der politischen Einstellungen. Dies trifft sowohl zu für vollständige Übereinstimmung der Parteipräferenz als auch für weitgehende Übereinstimmung (drei der vier Freunde teilen eine Einstellung).

Schließlich soll noch untersucht werden, ob Unterschiede festzustellen sind, je nachdem, welche politischen Präferenzen der Befragte selbst hat. Unabhängig, ob der Befragte Anhänger der CDU/CSU oder aber von SPD/FDP ist, nimmt mit der Dichte der Freundesbeziehung auch die Übereinstimmung der politischen Einstellungen zu. Für Befragte mit voll verbundenen Freundesnetzen gilt generell: Rund 60% der Netze sind nach Parteipräferenz voll übereinstimmend, weitere rund 30% sind überwiegend übereinstimmend, d.h. drei der vier Personen teilen eine Parteipräferenz.

Als drittes Merkmal der Befragten und der genannten Freunde soll auf die Angaben zur beruflichen Stellung für Befragte und Freunde eingegangen werden.

Für die von uns gewählte Fragestellung wollen wir im folgenden prüfen, ob der Befragte und dessen Freunde gemeinsam einer Gruppe gleicher beruflicher Stellung angehören, oder aber ob die Freunde einer Grup-

pe höherer oder niedrigerer beruflicher Stellung als der Befragte zuzurechnen sind. Die im ALLBUS sehr differenziert erfaßten Daten zur beruflichen Stellung wurden entsprechend rekodiert.

Tabelle 2: Berufliche Stellung von Befragten und Freund A, B, C

Typus der Freundes- beziehung	Befragter und A, B,C haben gleiche berufl.Stellung	Wahl nach oben A, B,C haben höhere berufl.Stellung als Befragter	Wahl nach unten A, B,C haben niedri- gere berufl.Stel- lung als Befragter
voll segmentiert	44,9%	20,7%	34,4%
überwiegend seg- mentiert	42,2%	25,9%	31,9%
überwiegend ver- bunden	41,8%	21,5%	36,7%
voll verbunden	45,0%	22,6%	32,4%

Über alle Typen von Freundesnetzen hinweg können wir feststellen, daß diese nach der Komposition der beruflichen Stellung zu rund 45% als voll übereinstimmend betrachtet werden können. Bei weiteren 21 bis 26% der Beziehungen ist die berufliche Stellung der Freunde eine höhere als die der Befragten. Für die verbleibenden Fälle ist die berufliche Stellung, der der Befragte angehört, eine höhere als diejenige seiner Freunde.

Im Gegensatz zu den Merkmalen Geschlecht und Parteipräferenz ist für berufliche Stellung daher kein ausgeprägter Effekt der Dichte der Freundesnetzwerke festzustellen.

Die Ergebnisse für die Merkmale Geschlecht, Parteipräferenz und berufliche Stellung wollen wir nachstehend zusammenfassen. Tabelle 3 zeigt, aufgegliedert nach der Dichte der Freundesbeziehungen, zu wieviel Prozent die jeweils vier Mitglieder der Freundesnetzwerke vollständig das gleiche Merkmal teilen.

Tabelle 3: Homogenität von Merkmalen nach Dichte der Beziehung

Freundesnetze nach Dichte der Beziehung		Volle Übereinstimmung von persönlichen Merkmalen zwischen den Befragten und deren Freunde A, B, C		
Typ	in % +)	Geschlecht %	Berufli.Stellung %	Parteipräferenz %
voll segmentiert	12,0	39,4	44,9	43,5
überwiegend segmentiert	25,5	29,8	42,2	41,0
überwiegend verbunden	8,9	31,6	41,8	40,6
voll verbunden	53,7	28,1	45,0	59,0

+) Anteil des Beziehungstyps an allen Fällen mit drei Freunden

Zusammenfassend kann festgestellt werden: Je stärker die Verbindung der Freundesnetze nur über den Befragten strukturiert ist (voll oder überwiegend segmentiert), desto größer der Anteil der Beziehungen, in denen alle Mitglieder dem gleichen Geschlecht angehören. Andererseits steigt mit der Dichte der Freundesnetze, d.h. der Anzahl der Beziehungen, die zwischen den genannten Freunden direkt bestehen, der Anteil der im Bezug auf das Merkmal Parteipräferenz homogenen Beziehungen. Für die berufliche Stellung findet sich für alle Beziehungen ein Effekt nahezu gleicher Stärke.

Für Freundesnetze unterschiedlicher Dichte sind also persönliche Merkmale der Netzwerkmitglieder unterschiedlich wichtig, wobei diese Unterschiede bei den beiden Extremtypen - dem voll segmentierten und dem voll verbundenen Freundesnetz - für Geschlecht und Parteipräferenz am prägnantesten ausfallen.

Für eng bzw. locker verbundene Netze (Bott 1957) sind diese persönlichen Merkmale als unterschiedlich wichtig zu betrachten. Da wir die Stärke der einzelnen Beziehung als gleich betrachten müssen, sind uns weitere Einblicke nicht möglich.

Dennoch kann zumindest vermutet werden, daß in voll segmentierten Netzen die Beziehung mehr durch Bedürfnisse - wie die Partizipation in unterschiedlichsten Bereichen - geprägt ist, die eher zu schwachen Verbindungen (Granovetter 1982) innerhalb des Netzes führen. Voll verbundene Netze können hingegen als durch starke Verbindungen geprägt angenommen werden, die am ehesten zu Personen bestehen und zu diesen aufrechterhalten werden können, die in wesentlichen Merkmalen einander ähnlich sind. Die politische Einstellung, repräsentiert durch Parteipräferenz, scheint ein solches Merkmal zu sein.

Die berufliche Stellung ist für alle Beziehungstypen ein offenbar gleichermaßen wichtiges Merkmal.

Wir wollen nun auf die zu Beginn formulierten Hypothesen zurückkommen und diese einem Test mit LISREL unterziehen. Daran anschließend soll überprüft werden, ob sich die Ergebnisse bedeutsam ändern, sofern die Dichte der Freundesbeziehung als zusätzlicher Faktor in die Analyse einbezogen wird, d.h. ein wichtiges Merkmal des sozialen Kontexts explizit zum Test der Stabilität der Hypothesen benutzt wird.

5. Spezifikation und Test des Strukturgleichungsmodells (LISREL-Modell)

Gehen wir nun auf die Spezifikation unseres Modells näher ein: LISREL-Modelle bestehen generell aus drei Komponenten (vgl. Jöreskog und Sörbom 1982), dem:

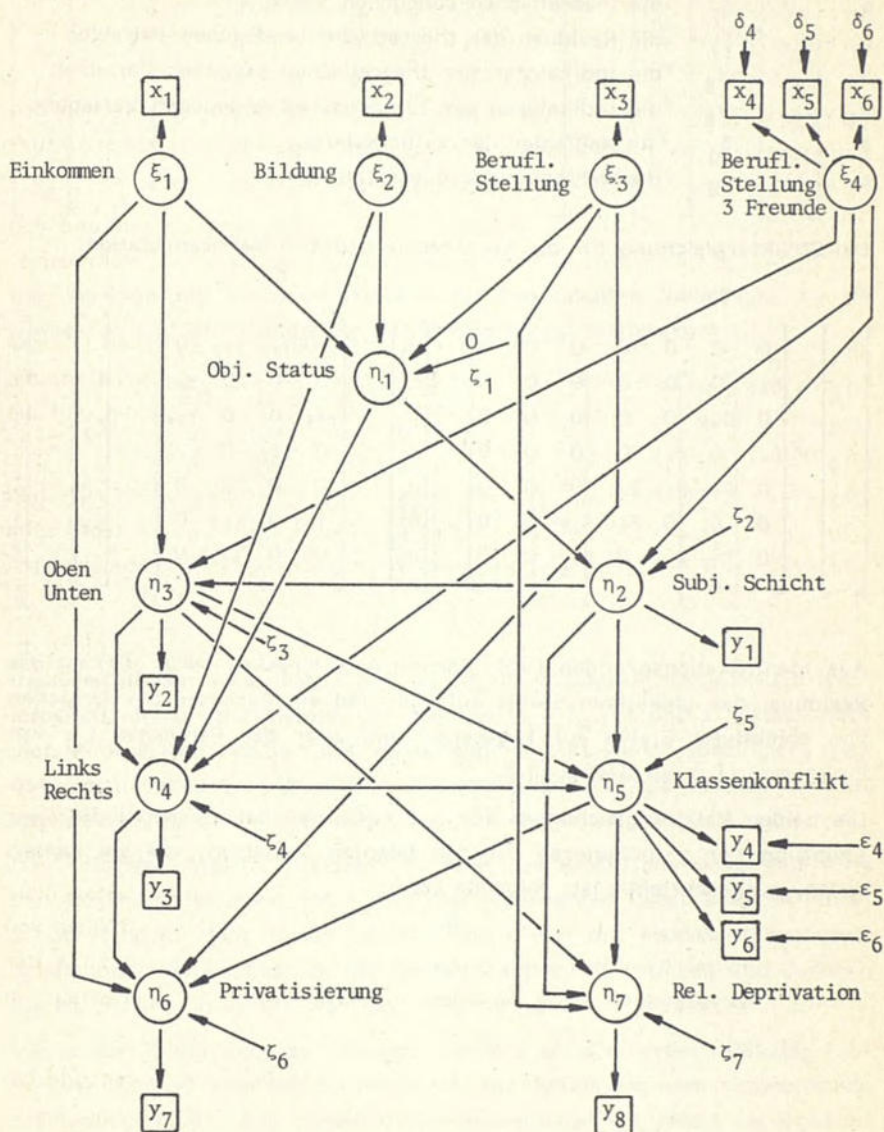
1. Strukturgleichungsmodell. In ihm werden die Beziehungen zwischen den latenten nicht direkt meßbaren Variablen spezifiziert. Das Strukturgleichungsmodell enthält die Kerntheorie.
2. Meßmodell. Es enthält die Beziehungen zwischen den latenten und den beobachteten Variablen. Das Meßmodell repräsentiert die Meßtheorie, d.h. die Korrespondenzhypothesen.
3. Annahmen, die sich auf das Struktur- und Meßmodell beziehen. Hierzu gehören u.a. die Annahmen über die Fehler und deren Kovariation, die Annahme der Multinormalverteilung der beobachteten Variablen und die Annahme über das Skalenniveau.

Unser Strukturgleichungsmodell mit den postulierten Beziehungen zwischen den latenten Variablen zeigt Abbildung 1. Eine ausführlichere technische Beschreibung des Modells einschließlich der notwendigen Steuerkarten findet sich bei Schmidt (1983).

Da die Variablen in dem Strukturgleichungsmodell latente, d.h. nicht direkt meßbare Variablen sind, müssen wir ein Meßmodell für die elf Konstrukte spezifizieren. Bei acht der elf Konstrukte steht nur ein Indikator pro Konstrukt zur Verfügung. Bei einem Konstrukt (objektiver Status) gibt es gar keinen direkt meßbaren Indikator, da nach unseren theoretischen Überlegungen die direkt meßbaren Einflußgrößen Bildung, Einkommen und berufliche Stellung, Determinanten und nicht Wirkungen des objektiven Status sind (vgl. Blalock 1982). Dies impliziert für diese neun Konstrukte, daß aufgrund von Identifikationsproblemen ein Meßfehler von Null angenommen werden muß. Mit anderen Worten wird eine 1 zu 1 Beziehung vorausgesetzt. Die Meßtheorie hierzu findet sich ebenfalls in Abbildung 1. Die Frageformulierungen finden sich in Anhang 1.

Somit kann das Ausmaß zufälliger und systematischer Meßfehler nur bei den Konstrukten berufliche Stellung der drei besten Freunde und Wahrnehmung von Klassenkonflikten ermittelt werden, denn bei diesen beiden Konstrukten stehen jeweils drei Indikatoren zur Messung zur Verfügung.

Abbildung 1: Pfaddiagramm des LISREL-Modells



In dem Modell sind:

$\xi_1 \dots \dots \xi_4$	die theoretischen exogenen Variablen
$\eta_1 \dots \dots \eta_7$	die theoretischen endogenen Variablen
$\zeta_1 \dots \dots \zeta_7$	die Residuen der theoretischen endogenen Variablen
$x_1 \dots \dots x_8$	die Indikatoren der theoretischen exogenen Variablen
$y_1 \dots \dots y_8$	die Indikatoren der theoretischen endogenen Variablen
$\delta_1 \dots \dots \delta_6$	die Meßfehler der x-Indikatoren
$\varepsilon_1 \dots \dots \varepsilon_8$	die Meßfehler der y-Indikatoren

Die Strukturgleichung für die Kerntheorie lautet in Matrizennotation:

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_{32} & 0 & 0 & 0 & 0 & 0 \\ \beta_{41} & 0 & \beta_{43} & 0 & 0 & 0 & 0 \\ 0 & \beta_{52} & \beta_{53} & \beta_{54} & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_{64} & \beta_{65} & 0 & 0 \\ 0 & \beta_{72} & \beta_{73} & 0 & \beta_{75} & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix} + \begin{bmatrix} 1.0 & \gamma_{12} & \gamma_{13} & 0 \\ 0 & 0 & 0 & \gamma_{24} \\ \gamma_{31} & 0 & 0 & \gamma_{34} \\ 0 & \gamma_{42} & 0 & \gamma_{44} \\ 0 & 0 & 0 & 0 \\ \gamma_{61} & 0 & \gamma_{63} & 0 \\ 0 & 0 & \gamma_{73} & 0 \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \\ \zeta_7 \end{bmatrix}$$

Aus Identifikationsgründen (vgl. Sörbom und Jöreskog 1980: 191) ist das Residuum des objektiven Status auf Null und ein Parameter der Ursachen von objektivem Status auf 1 gesetzt, und zwar der Parameter γ_{11} von Einkommen (ξ_1) auf Status (η_1).

Die beiden Matrixengleichungen für das Meßmodell haben, wobei die Verknüpfung der y-Indikatoren mit den latenten Variablen, die sie messen sollen, ausgeschrieben ist, folgende Form:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} 0 & 1.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{45} & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{65} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix} \cdot \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ 0 \\ 0 \end{bmatrix}$$

Die Verknüpfung der x-Indikatoren mit den latenten Variablen, die sie messen sollen, ist in folgender Matrizengleichung wiedergegeben:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & \lambda_{44} \\ 0 & 0 & 0 & \lambda_{54} \\ 0 & 0 & 0 & \lambda_{64} \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix}$$

Die Gleichungen entsprechen dem in dem Pfaddiagramm dargestellten Kausalmodell. Nach Spezifikation der Matrizen Λ_y , Λ_x , B und Γ müssen noch vier weitere Matrizen für das vollständige LISREL-Modell spezifiziert werden (vgl. hierzu Schmidt 1983). Die verwendete Korrelationsmatrix findet sich in Anhang 2.

Das Programm LISREL (Version V) testet das spezifische Modell mit Hilfe der Maximum-Likelihood (ML) oder der Unweighted Least Squares (ULS) Schätzmethode. Für unser Modell haben wir die Maximum-Likelihood-Schätzungen verwendet, da die Stichprobe ziemlich groß ist und dadurch die ML-Schätzung optimal im Sinne möglichst großer Präzision ist.

Wegen der Einführung der latenten Variable ohne direkten Indikator (objektiver Status) arbeitet die Prozedur zur Erzeugung von automatischen Startwerten nicht. Aus diesem Grund verwendeten wir selbst vorgegebene Startwerte.

Die Beurteilung der Modellgüte stellt einen besonders wichtigen Teil der Anwendung solch komplexer Verfahren dar. Sie umfaßt dabei drei Aspekte:

- a) Detaillierte Analyse der geschätzten Parameter
- b) Beurteilung der Maße für die Gesamtanpassung des Modells
- c) Beurteilung der Maße für Anpassung der einzelnen Parameter sowie für die Modellmodifikation.

Behandeln wir zunächst den in a) genannten Aspekt. Die zu prüfenden Schätzungen sind:

1. Parameterschätzungen
2. Standardschätzfehler
3. Quadrierte multiple Korrelationen der Variablen
4. Determinationskoeffizienten für die Gleichungen des Meß- und Strukturmodells
5. Korrelationen der Schätzungen.

In dem vorliegenden Modell hat keine der oben in 1-4 genannten Größen unplausible Werte in folgendem Sinn: Es tritt keine geschätzte Korrelation größer 1 auf, keine negative Varianz und keine extrem hohen Standardschätzfehler. Das einzige Problem stellen die hohen Korrelationen zwischen γ_{21} (objektiver Status auf subjektive Schicht) und γ_{12} sowie γ_{13} dar (Bildung und berufliche Stellung auf objektiven Status).

Die Korrelationen besagen, daß der Einfluß von Bildung und beruflicher Stellung auf die 'künstliche' Variable objektiver Status und dessen Einfluß auf subjektive Schichtidentifikation nur schwer trennbar sind. Dies liegt primär daran, daß objektiver Status aus Identifikationsgründen ohne Fehlerkomponente als Linearkombination der drei Größen Einkommen, Bildung und berufliche Stellung konstruiert wurde.

Gehen wir nun auf die Parameterschätzungen näher ein. Wegen der besseren Vergleichbarkeit konzentrieren wir uns auf die standardisierten Koeffizienten.

Der Übersichtlichkeit halber finden sich die standardisierten Parameter der Kern- und der Meßtheorie in dem Pfaddiagramm in Abbildung 2.

Behandeln wir jetzt die Interpretation der Koeffizienten. Zunächst muß festgestellt werden, daß alle Koeffizienten in Abbildung 2 die vorherge-

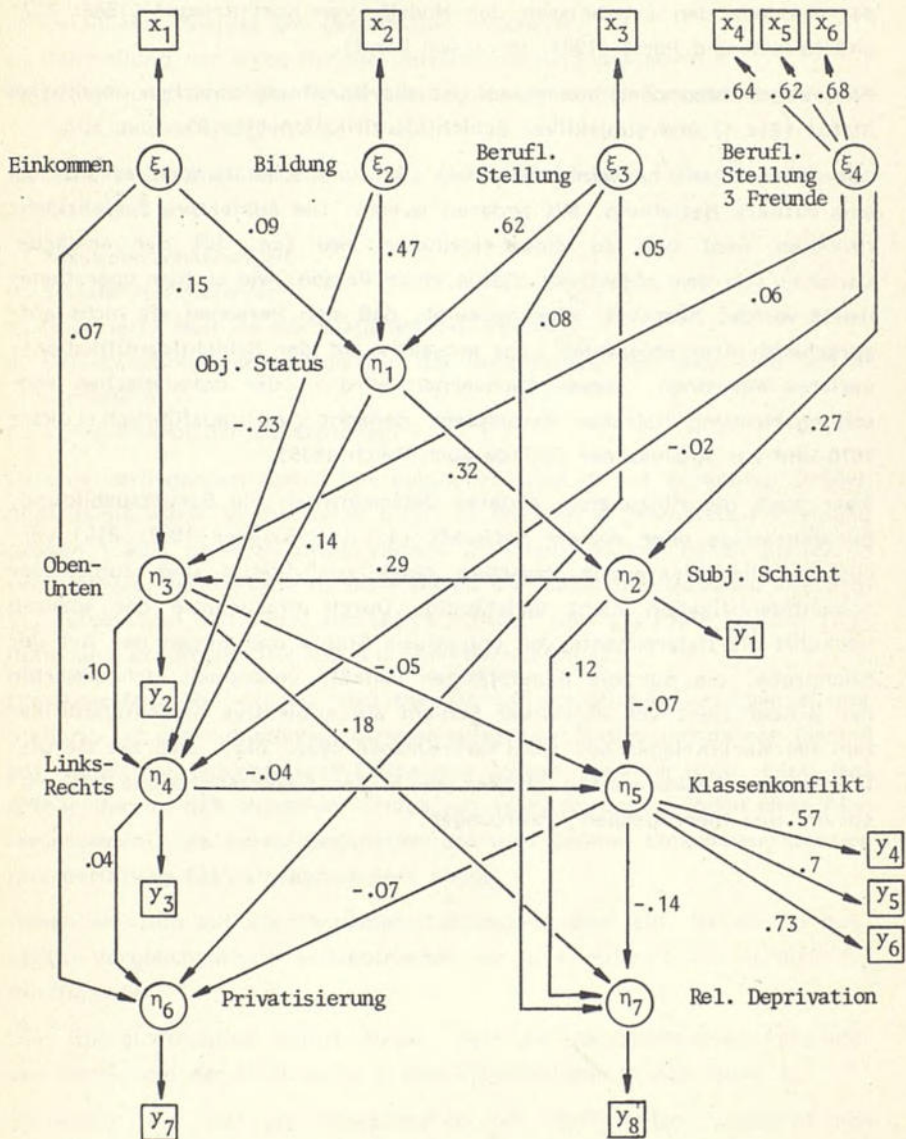
sagten Vorzeichen aufweisen. Objektiver Status wird quantitativ nur schwach von Einkommen (.09), dagegen stark von Schulbildung (.47) und noch stärker von beruflicher Stellung (.62) beeinflusst. Dies entspricht in der Tendenz den Ergebnissen der Modelle von Kort-Krieger (1982: 217) und Krauth und Porst (1983, in diesem Band).

Theoretisch besonders interessant ist die Beziehung zwischen objektivem Status (Eta 1) und subjektiver Schichtidentifikation (Eta 2).

Der β -Koeffizient hat einen Wert von .32, und somit handelt es sich um eine mittlere Beziehung. Mit anderen Worten: Die subjektive Schichtidentifikation wird nur zu einem kleineren Teil (ca. 10% der erklärten Varianz) von dem objektiven Status einer Person, wie er hier operationalisiert wurde, bestimmt. Dies bedeutet, daß sich Personen oft nicht entsprechend ihrer objektiven Lage subjektiv auf der Schichtidentifikationsvariable einordnen. Dieser Sachverhalt wird in der marxistischen Forschungsrichtung 'falsches Bewußtsein' genannt (vgl. ausführlich Lukacs 1970 und zur Analyse der Gründe auch Reich 1935).

Aber auch die Hinzunahme anderer Determinanten wie Berufsausbildung, Berufsprestige oder soziale Herkunft (s. Kort-Krieger 1982: 215) verringern die Diskrepanz zwischen objektivem Status und subjektiver Schichtidentifikation nicht vollständig. Durch Hinzunahme der sozialen Herkunft als Determinante von objektivem Status und wegen der Art der Stichprobe, die nur die Berufstätigen umfaßt, verdoppelt sich immerhin der β -Koeffizient von objektiver Schicht auf subjektive Schichtidentifikation bei Kort-Krieger auf .60 (Kort-Krieger 1982: 215). Daß bei Berufstätigen der Zusammenhang stärker als in der Gesamtstichprobe ist, entspricht den theoretischen Erwartungen.

Abbildung 2: Geschätzte Koeffizienten des LISREL-Modells



Auffällig ist der im Vergleich zum objektiven Status des Befragten fast gleich starke Einfluß der beruflichen Stellung der drei besten Freunde auf die subjektive Schichtidentifikation (.27). Damit ist der Effekt dieser Kontextvariablen fast genau so stark wie die Wirkung der Merkmale des Befragten selbst.

Auch wenn man in Rechnung zieht, daß der Effekt der beruflichen Stellung der Freunde durch das Meßmodell meßfehlerkorrigiert ist, während dies für die Merkmale der Befragten nicht gilt, bleibt dies eine bemerkenswerte Bestätigung der qualitativen Hypothesen von Marx über die Wirkung des sozialen Kontexts auf die Entwicklung des Klassenbewußtseins (Marx 1976: 117).

Damit stellt sich aber als weitere Frage, wie dieser Kontexteffekt zustande kommt. Neben Replikationen, um die Stabilität dieses Effekts in unabhängigen Stichproben zu prüfen, müssen Mikrotheorien angewandt werden, um diesen Effekt zu erklären (vgl. Lindenberg und Wippler 1978). Ohne eine genaue theoretische Rekonstruktion vorzunehmen, scheinen mindestens zwei Mikrotheorien Erklärungen für diesen Kontexteffekt zu liefern. Aus der Theorie sozialer Vergleichsprozesse (Festinger 1957) kann man Hypothesen entnehmen, die erklären, warum bei Ansteigen der beruflichen Stellung der Freunde auch die subjektive Schichtidentifikation der Befragten steigen könnte. Ebenso liefert die Theorie des Lernens am Modell (Bandura 1971) Hypothesen, die den gefundenen Effekt erklären könnten. Jeweils liegt die Idee zugrunde, daß erst über subjektive Nutzensvorstellungen der Kontexteffekt auf Verhalten wirkt. Explizit haben dies Blalock und Wilken (1979) herausgearbeitet, die argumentieren, daß Kontexteffekte nur indirekt über individuelle Bewertungen und Erwartungen Handeln bestimmen. Die Entscheidung der Befragten, sich einer Schicht zuzuordnen, wird von uns als Handeln interpretiert, und daher können die oben dargestellten Überlegungen auf die Erklärung der subjektiven Schichtzuordnung angewendet werden.

Die Links-Rechts-Einstufung einer Person wird erstaunlicherweise durch keine der kausal vorangehenden Variablen stark beeinflusst. Die erklärte Varianz beträgt nur .03. Subjektive Schichtidentifikation selbst wirkt mit .29 auf die Oben-Unten-Einschätzung. Dies zeigt quantitativ, daß die Oben-Unten-Skala nicht als Parallelmessung zur subjektiven Schichtidenti-

fikation angesehen werden kann. Aber auch die berufliche Stellung der drei Freunde hat einen wenn auch geringen Effekt (.06).

Je höher der objektive Status einer Person und desto mehr sich eine Person in der gesellschaftlichen Rangordnung oben einordnet, desto mehr ordnet sie sich rechts ein (.14 und .10). Der Effekt von Bildung ist hingegen entgegengesetzt, d.h. je höher die Bildung, desto eher orientieren sich die Personen links (-.23).

Die Wahrnehmung von Klassenkonflikten kann sogar noch etwas weniger als die Links-Rechts-Orientierung durch das spezifizierte Modell erklärt werden (erklärte Varianz ca. .01). Die Vorzeichen der allerdings sehr niedrigen Koeffizienten sind wie erwartet, d.h. je höher die subjektive Schichtzuordnung, die Oben-Unten-Einstufung und je mehr eine Person rechts orientiert ist, desto weniger stark werden Klassenkonflikte wahrgenommen. Auch hier muß überlegt werden, ob die Kerntheorie zumindest sehr unvollständig, die Meßtheorie nicht adäquat ist, z.B. weil Personen mit geringem politischen Interesse eher Zufallsantworten geben (vgl. Converse 1964) oder weitere Annahmen des Modells, z.B. die Linearitätsannahme, nicht adäquat sind. Ein ähnliches Phänomen zeigt sich bei dem Privatisierungsitem. Die Vorzeichen der Koeffizienten sind wie vorhergesagt, aber die erklärte Varianz nur .02.

Bezogen auf die Erklärung der relativen Deprivation zeigt sich folgendes: Je höher die berufliche Stellung (.08), je höher die subjektive Schichtidentifikation (.12), je höher die Oben-Unten-Einstufung (.18) und je weniger stark Klassenkonflikte wahrgenommen werden, desto eher glauben Personen, ihren gerechten Anteil zu bekommen (Relative Gratifikation). Auch hier aber ist die erklärte Varianz niedrig.

Es stellt sich die Frage, wie dies 'falsche Bewußtsein' zu erklären ist. Hier können wir auf drei Ebenen ansetzen. Zum einen kann die Kerntheorie unvollständig sein. Zum zweiten kann die Meßtheorie inadäquat und zum dritten können die weiteren Annahmen falsch sein. Eine differenzierte Analyse muß alle drei Ebenen für die Modellmodifikation berücksichtigen.

In dieser Arbeit sollen Alternativen zu der spezifizierten Kerntheorie benutzt werden, die die Merkmale des sozialen Netzwerks (der drei besten Freunde) berücksichtigen (vgl. Abschnitt 6).

Es könnte aber auch z.B. eine Annahme der Meßtheorie, und zwar die des Intervallskalenniveaus, fallengelassen werden und das gleiche Modell mit polychorischen bzw. polyseriellen Korrelationen gerechnet werden (vgl. hierzu Schmidt 1983).

Betrachten wir nun noch die Koeffizienten des Meßmodells. Die berufliche Stellung der drei besten Freunde hat eine zufriedenstellende formale Gültigkeit, was sich in den Faktorenladungen von .64, .62 und .68 manifestiert. Dabei zeigt sich, daß die Indikatoren nahezu gleiche formale Gültigkeit haben und daher als Parallelitems gelten können. Auch die Faktorenladungen der Klassenkonfliktitems sind als befriedigend zu bezeichnen. Allerdings fällt hier auf, daß die Ladungen stärker variieren.

Im folgenden geben wir in Form einer Tabelle nochmals eine kompakte Übersicht über die erklärten Varianzen der latenten und der beobachteten Variablen. Hierbei werden nur die Variablen aufgeführt, deren erklärte Varianz empirisch ermittelt und nicht a priori festgesetzt wurde.

Tabelle 4: Varianzerklärung der Variablen des Strukturmodells und des Meßmodells

a) Strukturmodell: latente endogene Variablen		b) Meßmodell: x und y Indikatoren	
	R ²		R ²
Subj. Schicht	.304	Beruf. Stellung A	.410
Oben-Unten	.167	Beruf. Stellung B	.382
Links-Rechts	.029	Beruf. Stellung C	.460
Klassenkonflikt	.013	Arbeitgeber-Arbeitnehmer	.319
Privatisierung	.019	Arm-Reich	.480
Relative Gratifikation	.116	Kapitalisten-Arbeiter	.537

Die Anpassungsmaße für das Gesamtmodell zeigen ohne Ausnahme befriedigende Werte und finden sich in Tabelle 5.

Tabelle 5: Anpassungsmaße für das Gesamtmodell

χ^2		62.68	
Freiheitsgrade		60	Likelihood-Ratio-Test
probability level		.381	inferenzstatistisch
(GFI)	Goodness of Fit Index	.991	deskriptive Maße der im Modell erklärten Varianz (GFI) und (AGFI) bzw. durchschnittlichen Residuenkorrelation dieses Modells (RMR).
(AGFI)	Adjusted Goodness of Fit Index	.985	
(RMR)	Root Mean Residual	.018	

AGFI ist ein Maß der im Modell erklärten Varianz, bei dem die Zahl der Freiheitsgrade berücksichtigt ist (s. Jöreskog und Sörbom 1982: 1.40).

Im Unterschied zum Likelihood-Ratio-Test ist das GFI- und AGFI-Maß unabhängig von der Stichprobengröße und relativ robust gegenüber der Verletzung der Multinormalverteilungsannahme.

Auch die Maße für die detaillierte Analyse der Anpassung des Modells, nämlich die Matrix der normalisierten Residuen und die Matrix der Modifikationsindices, zeigen bis auf eine Ausnahme Werte, die darauf schließen lassen, daß das Modell nicht mehr modifiziert zu werden braucht. Normalisierte Residuen, welche größer als 2 sind, wären ein Anzeichen für einen Spezifikationsfehler. Dies ist nur bei einem auf Null festgelegten Parameter der Fall. Da dieser jedoch nur unwesentlich über 2 liegt und der entsprechende Modifikationsindex unter der von Jöreskog und Sörbom angegebenen Grenze von 5 liegt, halten wir eine Modifikation nicht für notwendig.

Um einen umfassenden Einblick in die Wirkungsketten komplexer Kausalmodelle zu erhalten, ist die Betrachtung der direkten Effekte nicht ausreichend. Man benötigt auch die Kenntnis der indirekten und totalen kausalen Effekte einer Variable in einem Modell auf die kausal nachfolgenden Variablen (vgl. Graff und Schmidt 1982).

Die totalen kausalen Effekte der vier exogenen latenten Variablen seien hierfür exemplarisch in Tabelle 6 wiedergegeben:

Tabelle 6: Totale und direkte kausale Effekte von ... (KSI) auf ETA

	KSI 1		KSI 2		KSI 3		KSI 4	
	Einkommen		Bildung		Berufliche Stellung		Berufl.Stellung der drei Freunde	
	direkt	total	direkt	total	direkt	total	direkt	total
ETA 1 Objektiver Status	.089	.089	.474	.474	.617	.617	0	0
ETA 2 Subjektiver Status	0	.028	0	.151	0	.197	.271	.271
ETA 3 Oben-Unten	.154	.162	0	.044	0	.057	.064	.143
ETA 4 Links-Rechts	0	.029	-.230	-.159	0	.093	-.023	-.009
ETA 5 Klassenkonflikt	0	-.012	0	-.006	0	-.021	0	-.026
ETA 6 Privatisierung	.070	.073	0	-.007	.046	.051	0	.001
ETA 7 Relative Gratifikation	0	.035	0	.027	.084	.121	0	.062

Am interessantesten ist hier der Vergleich des direkten und des totalen kausalen Effekts von Bildung auf die Links-Rechts-Einstufung. Direkt hat Bildung einen Effekt in dem Sinne, daß sich Personen mit steigender Bildung eher links einordnen. Dieser Effekt wird jedoch durch den indirekten Effekt über objektiven Status wieder teilweise aufgehoben.

Steigende Bildung führt demnach zu steigendem objektiven Status, der wiederum führt aber zu einem Anwachsen der Tendenz, sich rechts einzustufen. Hier zeigt erst die Zerlegung des Gesamteffekts, daß sich direkte und indirekte Effekte wegen des unterschiedlichen Vorzeichens teilweise aufheben.

6. Der Einfluß der Segmentierung der Freundschaftsbeziehung auf die Variablen des LISREL-Modells

In diesem Teil wollen wir überprüfen, inwieweit sich bei den in Abschnitt 5 behandelten Variablen Mittelwerte, Korrelationen und/oder die Koeffizienten des LISREL-Modells verändern, wenn wir die Art, d.h. die Konnektivität der Freundschaftsbeziehung berücksichtigen. Als weitere Variable beziehen wir die Parteisympathie für die SPD ein. Die Struktur der Freundschaftsbeziehungen kann ausgehend von zwei Extremsituationen betrachtet werden (vgl. Teil 4, Tabelle 1).

Bei segmentierten Freundesnetzen läuft der Kontakt ausschließlich durch und über den Befragten ab. Hingegen sind voll verbundene Netze dadurch charakterisiert, daß nicht nur der Befragte mit allen Freunden verbunden ist, sondern sich auch diese Freunde direkt untereinander gut kennen oder befreundet sind.

Ziel dieses Abschnitts ist es, die in Teil 4 dargestellten Überlegungen zur Struktur der Freundschaftsbeziehungen mit dem im Abschnitt 5 wiedergegebenen LISREL-Modell zu verknüpfen. Mit anderen Worten: Es soll der Einfluß zunehmender Segmentierung bzw. Verbundenheit der Freundschaftsnetze auf die Mittelwerte, Streuungen, Korrelationen und Parameter der Kern- und Meßtheorie der Variablen des LISREL-Modells näher untersucht werden.

Für den Vergleich haben die genannten Größen jeweils unterschiedliche Bedeutung. Die Mittelwerte, Streuungen und Korrelationen stellen voneinander unabhängige deskriptive Kenngrößen dar. Hingegen ist der Vergleich der Parameter der Kern- und der Meßtheorie des LISREL-Modells in den Gruppen für die multivariate Prüfung der Hypothese der Invarianz des Modells über die Gruppen hinweg von Bedeutung.

Mit dem Vergleich der Parameter der Kerntheorie soll die Invarianz der in Abschnitt 2 aufgestellten Hypothesen über die unterschiedlichen Typen von Freundschaftsnetzen geprüft werden.

In einem ersten Schritt vergleichen wir nun die Mittelwerte und die Streuungen in den vier Gruppen. Daran anschließend stellen wir ausgewählte Korrelationen der zwei Extremgruppen dar. Den Abschluß bildet der Vergleich ausgewählter Koeffizienten des LISREL-Modells.

Tabelle 7 enthält die Mittelwerte und die Streuungen für die vier Typen von Freundschaftsnetzen. Bezüglich der inhaltlichen Bedeutung der Mittelwerte und der Streuungen verweisen wir auf die Codierung und die Wertebereiche der Variablen, die sich im Anhang befinden.

Bei den demographischen Merkmalen ist auffällig, daß mit ansteigender Segmentierung der Gruppen tendenziell auch die Mittelwerte der demographischen Variablen wie Einkommen, Schulabschluß und berufliche Stellung steigen. Auch die subjektive Schichtzuordnung und die Oben-Unten-Einordnung sind bei der segmentierten Gruppe am höchsten.

Bei der Links-Rechts-Einordnung ist die Tendenz zur Mitte der überwiegend segmentierten Gruppe und der überwiegend verbundenen Gruppe, im Vergleich zu den beiden Extremgruppen, plausibel. In diesen beiden Gruppen besteht einerseits mehr Heterogenität als in der voll segmentierten Gruppe, andererseits dürfte aber weniger Uniformität als in der voll verbundenen Gruppe vorliegen.

Bei den Konfliktvariablen, dem Privatisierungsitem und der Frage nach dem gerechten Anteil ist das Bild bei der voll segmentierten Gruppe am eindeutigsten. Sie weist die stärkste Wahrnehmung von Konflikten auf und glaubt am wenigsten, einen gerechten Anteil zu bekommen. Hingegen ist die Sympathie für die SPD in der überwiegend verbundenen Freundschaftsgruppe am höchsten.

Betrachten wir nun einige ausgewählte Korrelationen in den zwei Extremgruppen (vgl. Tabelle 8). Hierbei ist allerdings zu beachten, daß die Unterschiede in geringem Umfang auch auf Unterschiede in den Streuungen zurückführbar sein können.

Tabelle 7: Vergleich der Mittelwerte und Streuungen in den Gruppen

	Gruppen											
	voll segmentiert			überwiegend segmentiert			überwiegend verbunden			voll verbunden		
	\bar{X}	s	N	\bar{X}	s	N	\bar{X}	s	N	\bar{X}	s	N
Einkommen	1705	1056	141	1671	1004	312	1690	1053	107	1546	894	624
Schulabschluß	2.97	1.14	211	2.76	1.08	453	2.77	1.08	158	2.55	.98	952
Berufliche Stellung	2.96	1.06	108	2.86	1.04	231	2.82	1.05	74	2.69	.98	466
Berufl. Stellung Freund A	2.88	1.09	172	2.68	1.06	364	2.63	1.05	129	2.51	1.02	768
Berufl. Stellung Freund B	2.84	1.21	155	2.58	1.08	351	2.65	1.00	121	2.44	.91	744
Berufl. Stellung Freund C	2.84	1.06	170	2.74	1.07	354	2.53	1.06	127	2.52	1.00	711
Subjektive Schicht	2.99	.68	205	2.96	.66	438	2.94	.71	156	2.83	.72	926
Oben-Unten	5.95	1.58	208	5.90	1.49	441	5.85	1.41	156	5.86	1.47	933
Links-Rechts	6.00	2.03	207	5.88	1.94	445	5.57	1.96	157	5.79	2.00	926
Konflikt Arbeitgeber-Arbeitnehmer	3.03	.66	209	2.97	.76	444	2.89	.81	152	2.88	.79	919
Konflikt Arm-Reich	3.20	.77	210	3.07	.83	434	2.96	.83	153	3.02	.89	926
Konflikt Kapital-Arbeiter	3.17	.82	206	3.17	.78	428	3.11	.84	147	3.14	.85	895
Privatisierung Post	3.66	2.15	211	3.53	2.17	449	3.65	2.02	158	3.64	2.17	946
Gerechter Anteil	2.82	.54	205	2.78	.60	427	2.70	.70	148	2.72	.61	910
Thermometer SPD	7.56	2.55	209	7.46	2.66	439	7.76	2.43	158	7.43	2.78	935

Tabelle 8: Vergleich der Korrelationen in den zwei Extremgruppen

Variablenpaare	Gruppe voll segmentiert	Gruppe voll verbunden
Einkommen - Subjektive Schicht	.15	.23
Bildung - Subjektive Schicht	.34	.42
Berufliche Stellung - Subjektive Schicht	.34	.44
Berufliche Stellung Freund A - Subjektive Schicht	.25	.33
Berufliche Stellung Freund B - Subjektive Schicht	.37	.33
Berufliche Stellung Freund C - Subjektive Schicht	.31	.35
Subjektive Schicht - Oben-Unten	.39	.32
Links-Rechts - SPD-Sympathie	-.25	-.38

Am deutlichsten ist die Beziehung zwischen der Links-Rechts-Orientierung und SPD-Sympathie. Je verbundener das Netzwerk der Freunde ist, desto stärker wird - zumindest deskriptiv - auch die Beziehung zwischen der Links-Rechts-Orientierung und SPD-Sympathie. Dies bestätigt zumindest indirekt die von Marx formulierte Hypothese aus dem 18. Brumaire über den Zusammenhang zwischen Kommunikation und Integration einerseits und Bewußtsein über politische Zustände andererseits.

Auffällig ist weiterhin die Zunahme des Einflusses der objektiven demographischen Merkmale auf die subjektive Schichtidentifikation mit zunehmender Verbundenheit des Netzwerks. Allerdings ist dieser nicht vollkommen konsistent, da der Zusammenhang in einem Fall absinkt.

Entgegen unserer Hypothese ist die Korrelation zwischen subjektiver Schicht und der Oben-Unten-Einschätzung in der voll segmentierten Gruppe höher. Dieser Unterschied dürfte auch nicht auf variierende Streuungen zurückzuführen sein, da aus Tabelle 7 ersichtlich ist, daß die Streuungen der beiden Variablen zwischen den Gruppen nicht sehr unterschiedlich sind.

Betrachten wir nun die Parameter der Meßtheorie und die für unsere Fragestellung zentralen Parameter der Kerntheorie in den zwei Extremgruppen. Da die Varianz-Kovarianzmatrix der Gruppen bereits sehr unterschiedlich ist, sehen wir hier von einer formalen Testung auf Gleichheit der Kovarianzen ab (vgl. Jöreskog und Sörbom 1982: V.2). Vielmehr vergleichen wir die geschätzten standardisierten Koeffizienten in den beiden Gruppen.

Tabelle 9 enthält die Parameter der Meßtheorie. Alle Koeffizienten sind mindestens auf dem 5%-Niveau signifikant. Die Koeffizienten der Meßtheorie variieren zwischen den Gruppen kaum bedeutsam. Bei der beruflichen Stellung der Freunde A, B, C fällt nur die stärkere Faktorenladung von .797 in der voll segmentierten Gruppe, verglichen mit .619 in der voll verbundenen Gruppe, auf.

Tabelle 9: Vergleich der Koeffizienten des Meßmodells in den Extremgruppen

Meßtheorie		voll segmentiert	voll verbunden
Berufliche Stellung } Freunde	Freund A Freund B Freund C	.664 .797 .665	.669 .619 .742
Klassen- kon- } flikt	Arbeitgeber-Arbeitnehmer Arm - Reich Kapitalisten-Arbeiter	.458 .579 .856	.518 .654 .768

Bei den Items zur Messung der Wahrnehmung von Klassenkonflikten sind die Faktorenladungen bei den Items Arbeitgeber-Arbeitnehmer und Arm-Reich in der voll verbundenen Gruppe höher als in der voll segmentierten Gruppe. Nur das am 'ideologischsten' formulierte Item 'Kapitalisten-Arbeiter' fällt in der voll segmentierten Gruppe, mit .856 auf der latenten Variablen, aus dem Rahmen.

Während also zwischen den Gruppen die Unterschiede in der Meßtheorie quantitativ relativ gering sind, ändert sich das Bild bei der Kerntheorie.

Hingegen muß jedoch ein technisches Problem bei der Interpretation berücksichtigt werden: Besonders die Koeffizienten der demographischen Merkmale, wie auch die aus ihnen gebildete neue latente Variable Status, sind mit Vorsicht zu interpretieren, da teilweise hohe Multikollinearität vorliegt.

Vergleichen wir nun die standardisierten partialisierten Koeffizienten (β , γ) der Strukturtheorie aus Tabelle 10 mit den Korrelationen aus Tabelle 8.

Bis auf eine Ausnahme, den Effekt von Status auf subjektive Schicht, sind alle wiedergegebenen Koeffizienten mindestens auf dem 5%-Niveau signifikant.

Tabelle 10: Vergleich ausgewählter Korrelationen und Parameter der Kerntheorie (β , γ) in den zwei Extremgruppen

Variablenpaare	Gruppe voll segmentiert		Gruppe voll verbunden	
	Korrelation γ bzw. β		Korrelation γ bzw. β	
Einkommen - Subjektive Schicht ¹	.15	.127	.23	.295
Bildung - Subjektive Schicht	.34		.42	
Beruf. Stellung - Subjektive Schicht	.34		.44	
Beruf. Stellung Freund A - Subjektive Schicht ²	.25	.371	.33	.289
Beruf. Stellung Freund B - Subjektive Schicht	.37		.33	
Beruf. Stellung Freund C - Subjektive Schicht	.31		.35	
Subjektive Schicht - Oben-Unten	.39	.36	.32	.23
Links-Rechts - SPD-Sympathie	-.25	-.26	-.38	-.38

¹Im LISREL-Modell wurde statt Einkommen, Ausbildung und Berufliche Stellung die latente Variable Status verwendet.

²Im LISREL-Modell wurde statt Berufliche Stellung Freund A, B, C die latente Variable Berufliche Stellung Freunde verwendet.

Aus Tabelle 10 ersieht man, daß die Effekte der sozialstrukturellen Merkmale Einkommen, Bildung und berufliche Stellung auf die subjektive Schichtidentifikation auch im LISREL-Modell (vermittelt über die latente Variable Status) bei der voll verbundenen Gruppe größer als bei der segmentierten Gruppe sind. Dies gilt sowohl für die Korrelationen als auch die partialisierten Koeffizienten des LISREL-Modells.

Bei der Wirkung der beruflichen Stellung der Freunde ändert sich hingegen das Bild. In der voll segmentierten Gruppe ist zumindest der Parameter des LISREL-Modells (.371) höher als der entsprechende Parameter in der voll verbundenen Gruppe (.289). Dies widerspricht unseren theoretischen Erwartungen, da wir für die voll verbundene Gruppe den stärksten Effekt erwartet hatten.

Bei der Beziehung zwischen subjektiver Schicht und Oben-Unten sowie Links-Rechts und SPD-Sympathie ändert sich gegenüber den Korrelationen fast nichts, so daß wir hierauf nicht näher einzugehen brauchen.

7. Zusammenfassung und Ausblick

In dieser Arbeit wurden sozialstrukturelle und individuelle Determinanten von subjektiver Schichtidentifikation und politischen Einstellungen untersucht. Einen ersten Schwerpunkt bildete dabei die Beschreibung und Wirkung der primären Umwelt des Befragten, operationalisiert durch die Erfassung der drei besten Freunde, einige ihrer demographischen Merkmale und ihre Beziehungen untereinander. Als Ergebnis der deskriptiven empirischen Analysen zeigte sich, daß mit Zunahme der Segmentierung der Freundschaftsnetze der Anteil der Beziehungen größer wurde, in denen alle Mitglieder dem gleichen Geschlecht angehören. Andererseits steigt mit der Dichte der Freundesnetze der Anteil der bezüglich der Parteipräferenz homogenen Beziehungen. Für die berufliche Stellung findet sich für alle Beziehungen ein Effekt nahezu gleicher Stärke.

Im zweiten Teil wurde ein komplexes Strukturgleichungsmodell spezifiziert und getestet, bei dem die berufliche Stellung der drei besten Freunde als Kontextvariable mit drei Indikatoren berücksichtigt wurde. Durch diese Vorgehensweise sollte multivariat und soweit möglich mit multiplen Indika-

toren der quantitative Effekt der Variablen aufeinander überprüft werden. Die Beziehungen zwischen den Einstellungen waren quantitativ meist sehr gering, was gegen das Vorhandensein geschlossener politischer Vorstellungen wie z.B. Linksorientierung und starke Wahrnehmung von Klassenkonflikten in der Gesamtbevölkerung spricht. Andererseits konnte ein starker Kontexteffekt der beruflichen Stellung der drei besten Freunde auf die subjektive Schichtidentifikation des Befragten nachgewiesen werden. Mit anderen Worten: Die subjektive Schichtzuordnung des Befragten nimmt mit steigender beruflicher Stellung der Freunde zu, auch wenn der objektive Status des Befragten selbst (determiniert durch Einkommen, berufliche Stellung und Bildung) konstant gehalten wird.

Im letzten Teil wurde geprüft, ob insbesondere der Zusammenhang zwischen objektivem und subjektivem Status (Klasse an sich und Klasse für sich), zwischen beruflicher Stellung der Freunde und subjektiver Schichtidentifikation sowie zwischen den politischen Einstellungen des Befragten wie z.B. Linksorientierung und Sympathie für die SPD in segmentierten Freundesnetzen geringer ist als in vollverbundenen Netzen. Auf Grund der Korrelationen konnte dies deskriptiv bestätigt werden. Auch im LISREL-Modell wurden diese Zusammenhänge im wesentlichen bestätigt. Damit konnte, wenn auch nur in einem ersten Versuch, die Fruchtbarkeit der Anwendung von Netzwerkdaten und -hypothesen gezeigt werden. Es wäre nun sinnvoll, die Gruppenvergleiche noch wesentlich auszudehnen. So wäre z.B. zu prüfen, ob nicht in homogenen und vollverbundenen Netzen noch wesentlich stärkere Kontexteffekte zu verzeichnen sind.

In diesem Sinn versteht sich die vorliegende Arbeit als Anstoß für weitere Analysen mit Netzwerkdaten.

Literatur

- Bandura, Albert, 1971: Vicarious and Self-Reinforcement Processes. S.228-278 in: Robert Glaser (Hrsg.), *The Nature of Reinforcement*, New York: Academic Press.
- Bernard, Russel H., Peter D. Killworth und Leonard Sailer, 1981: A Review of Informant Accuracy in Social Network Data. S.153-186 in: Hans J. Hummell und Wolfgang Sodeur (Hrsg.), *Modelle für Ausbreitungsprozesse in sozialen Strukturen*. Duisburg: Verlag der Sozialwissenschaftlichen Kooperative.
- Blalock, Hubert M. Jr. und Paul H. Wilken, 1979: *Intergroup Processes. A Micro-Macro Perspective*. New York: Free Press.
- Blalock, Hubert M. Jr., 1982: *Conceptualization and Measurement in the Social Sciences*. Beverly Hills: Sage.
- Bott, Elizabeth, 1957: *Family and Social Network*. London: Tavistock.
- Boyd, Lawrence R. Jr. und Gudmund R. Iversen, 1979: *Contextual Analysis: Concepts and Techniques*. Belmont, Cal.: Wadsworth.
- Centers, Richard, 1949: *The Psychology of Social Classes*. New York: Russell and Russell.
- Clar, Michael, 1981: *Methodologische Probleme der Mehrebenenanalyse*. Diss. Hamburg.
- Converse, Philip E., 1964: The Nature of Belief Systems in Mass Publics. S.75-169 in: David E. Apter (Hrsg.), *Ideology and Discontent*. Glencoe, Ill: Free Press.
- Erbring, Lutz und Alice Young, 1980: Individuals and Social Structure. Contextual Effects as Endogeneous Feedback. S.25-95 in: Edgar F. Borgatta und David J. Jackson (Hrsg.), *Aggregate Data. Analysis and Interpretation*. Beverly Hills: Sage.
- Esser, Hartmut, 1982: On the Explanation of Contextual Effects on Individual Behavior: The Case of Language Acquisition by Migrant Workers. S.131-165 in: Werner Raub (Hrsg.), *Theoretical Models and Empirical Analyses. Contributions to the Explanation of Individual Actions and Collective Phenomena*. Utrecht: Explanatory Sociology Publications.
- Festinger, Leon, 1957: *A Theory of Cognitive Dissonance*. New York: Harper and Row.
- Fischer, Claude, 1982: *To Dwell among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.
- Gorz, André, 1971: *Abschied vom Proletariat*. Frankfurt: Europäische Verlagsanstalt.
- Graff, Jörg und Peter Schmidt, 1982: A General Model for Decomposition of Effects. S.131-148 in: Karl G. Jöreskog und Herman Wold (Hrsg.), *Systems under Indirect Observation: Causality, Structure, Prediction*. Amsterdam: North Holland.
- Granovetter, Mark, 1982: *Alientation Reconsidered: The Strength of Weak Ties*. Connections, Vol. V, Nr.2.

- Hummell, Hans J., 1974: Probleme der Mehrebenenanalyse. Stuttgart: Teubner.
- Jackman, Mary R. und Robert W. Jackman, 1973: An Interpretation of the Relation between Objective and Subjective Social Status. *American Sociological Review* 38: 569-582.
- Jöreskog, Karl G. und Dag Sörbom, 1982: LISREL Version V, User's Guide. Chicago: International Educational Services.
- Katz, Elihu und Paul F. Lazarsfeld, 1962: Persönlicher Einfluß und Meinungsbildung. München: R. Oldenbourg.
- Klingemann, Hans D., 1982: What 'Left' and 'Right' Means to Mass Publics: Variations in the Understanding of Political Symbols. Mimeo (paper read at the XIth World Congress of the International Political Sciences Associations, Rio de Janeiro, Brazil).
- Kort-Krieger, Ute, 1982: Structural Determinants of Objective and Subjective Status. S.206-226 in: Werner Raub (Hrsg.), *Theoretical Models and Empirical Analyses*. Utrecht: Explanatory Sociology Publications.
- Krauth, Cornelia und Rolf Porst, 1983: Sozioökonomische Determinanten von Einstellungen zu Gastarbeitern (in diesem Band).
- Laumann, Edward O., 1973: Bonds of Pluralism: The Form and Substance of Urban Social Networks. New York: Wiley.
- Laumann, Edward O. und Franz U. Pappi, 1976: Networks of Collective Action. New York: Academic Press.
- Lindenberg, Siegwart und Reinhard Wippler, 1978: Theorienvergleich: Elemente der Rekonstruktion. S.219-231 in: Karl O. Hondrich und Joachim Matthes (Hrsg.), *Theorienvergleich in den Sozialwissenschaften*. Darmstadt: Luchterhand.
- Lindner, Claus, 1972: Theorie der Revolution. München: Goldmann.
- Lukacs, Georg, 1970: Geschichte und Klassenbewußtsein. Neuwied: Luchterhand.
- Marx, Karl, 1976: Der 18. Brumaire des Louis Bonaparte. Berlin.
- McCallister, Lynne und Claude Fischer, 1978: A Procedure for Surveying Personal Networks. *Sociological Methods and Research*, 7: 131-148.
- Müller, Walter und Karl U. Mayer, 1980: Klassenlagen und soziale Schichtung in der Gesellschaft der Bundesrepublik Deutschland. Referat zum 1. Deutsch-Sowjetischen Soziologen-Symposium, Köln, 7.-8. Oktober 1980.
- Pappi, Franz U., 1973: Sozialstruktur und soziale Schichtung in einer Kleinstadt mit heterogener Bevölkerung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Heft 1: 23-74.
- Pappi, Franz U. und Gunter Wolf, 1981: Angaben über beste Freunde im Interview: Zuverlässigkeit und Gültigkeit. Kiel: unveröffentlichtes Manuskript.
- Reich, Wilhelm, 1968: Was ist Klassenbewußtsein? Ein Beitrag zur Neuformierung der Arbeiterbewegung. Amsterdam: Verlag de Munter.

- Scheuch, Erwin K., 1965: Die Sichtbarkeit politischer Einstellungen im alltäglichen Verhalten. S.169-214 in: Erwin K. Scheuch und Rudolf Wildenmann (Hrsg.), Zur Soziologie der Wahl. Köln und Opladen: Westdeutscher Verlag.
- Schmidt, Peter, 1982: Subjective Status Identification, Class Consciousness and Political Attitudes. S.227-252 in: Werner Raub (Hrsg.), Theoretical Models and Empirical Analyses. Utrecht: Explanatory Sociology Publications.
- Schmidt, Peter, 1983: LISREL V - ein Programm zur Analyse komplexer Kausalstrukturen bei experimentellen und nichtexperimentellen Forschungsdesigns. S.103-154 in: Helmut Wilke u.a. (Hrsg.), Statistik Software in der Sozialforschung. Berlin: Quorum.
- Sörbom, Dag und Karl G. Jöreskog, 1980: The Use of LISREL in Sociological Model Building. S.179-199 in: Edgar F. Borgatta und David J. Jackson (Hrsg.), Factor Analysis and Measurement in Sociological Research. Beverly Hills: Sage.
- Ziegler, Rolf, 1982: Die Struktur von Freundschafts- und Verwandtschaftsbeziehungen. München, unveröffentlichtes Manuskript.

Anhang 1: Fragebogensauszug

Erfassung der Merkmale der drei besten Freunde des Befragten (Split B)

<p>Wir haben jetzt einige Fragen zu Ihrem Bekanntenkreis: Denken Sie bitte an die drei Personen, mit denen Sie näher befreundet sind und mit denen Sie sich am häufigsten treffen. Es kann sich dabei sowohl um Verwandte als auch um Nicht-Verwandte handeln, nur nicht um Personen, die mit Ihnen im selben Haushalt wohnen. Ich möchte gerne einige statistische Angaben über jeden Einzelnen haben, wie z.B. Beruf, Alter usw. Um es ein wenig einfacher zu machen: können Sie mir bitte für alle drei Personen den Vornamen angeben:</p> <p><u>INT.:</u> Vornamen jeweils unter den Buchstaben A/ B/ C eintragen und "männlich" oder "weiblich" einkreisen. Die Vornamen bitte auch in den folgenden Fragen zur Kennzeichnung der drei Personen verwenden.</p> <p>Falls überhaupt keine Personen als Freunde oder Bekannte genannt → weiter mit Frage 16.</p> <p>Zuerst für "A" die Fragen 9 - 14 stellen, dann für "B", dann für "C" und Antworten unter dem entsprechenden Buchstaben eintragen.</p>			
<p><u>INT.:</u> Vornamen eintragen →</p>			
	Person A	Person B	Person C
9	keinen Freund bzw. keinen zweiten oder dritten Freund angegeben	0	0
	männlich	1	1
	weiblich	2	2
10	Wie alt ist A (B/ C)?	15/16	26/27
	<u>INT.:</u> Falls "weiß nicht", schätzen lassen	Jahre	Jahre
11	Ist irgendeine der eben genannten Personen mit Ihnen verwandt?	17/18	28/29
	<u>INT.:</u> weisse Liste 2 vorlegen	Kennziffer: eintr.	Kennziffer eintr.
	Sagen Sie mir bitte, was auf dieser Liste für A (B/ C) zutrifft.		
	<u>INT.:</u> Kennziffer eintragen		
12	Sagen Sie mir bitte nun, ob A (B/ C) erwerbstätig ist oder was sonst auf dieser Liste auf ihn/ sie zutrifft.	19	30
	<u>INT.:</u> rosa Liste 3 vorlegen, Kennziffer eintragen	Kennziffer eintr.	Kennziffer eintr.
	weiß nicht . . . 8	weiß nicht . . . 8	weiß nicht . . . 8
13	Bitte ordnen Sie die derzeitige oder die letzte berufliche Stellung von A (B/ C) nach dieser Liste ein.	20/21	31/32
	<u>INT.:</u> graue Liste 4 vorlegen, Kennziffer eintragen	Kennziffer eintr.	Kennziffer eintr.
	Beruf nicht bekannt . . . 98	Beruf nicht bekannt . . . 98	Beruf nicht bekannt . . . 98
	noch nie berufstätig . . . 96	noch nie berufstätig . . . 96	noch nie berufstätig . . . 96
14	Können Sie mir sagen, welche Partei A (B/ C) gewöhnlich wählt?	22	33
	<u>INT.:</u> Falls "andere Partei", nachfragen, um welche es sich handelt	CDU/ CSU 1	CDU/ CSU 1
		SPD 2	SPD 2
		FDP 3	FDP 3
		andere Partei, welche?	andere Partei, welche?
		4	4
	eintragen	eintragen	eintragen
	nicht wahlberechtigt . . . 6	nicht wahlberechtigt . . . 6	nicht wahlberechtigt . . . 6
	verweigert . . . 7	verweigert . . . 7	verweigert . . . 7
	weiß nicht . . . 8	weiß nicht . . . 8	weiß nicht . . . 8
	weiter mit "B", P9	weiter mit "C", P9	weiter mit "C", P9
	Falls keine zweite Person genannt, weiter mit Frage 16	Falls keine zweite Person genannt, weiter mit Frage 16	Falls keine zweite Person genannt, weiter mit Frage 16

15	Sagen Sie mir bitte nun abschließend zu diesem Thema noch, ob auch A und B miteinander befreundet sind, oder ob sie nicht miteinander befreundet sind.			
	<i>INT.:</i> Antwort bei A+B einkreisen. Falls 3 Personen genannt, weiterfragen: Und wie ist es bei A und C? Und wie ist es bei B und C?			
		befreundet	nicht befreundet	keine 3. Person genannt
	A und B	1	2	
	A und C	1	2	0
	B und C	1	2	0
				45
				46
				47
				9

Version A: '... Personen, mit denen Sie häufig privat zusammen sind';
 '... drei Personen, mit denen Sie am häufigsten privat zusammen sind'
 statt '... Bekanntenkreis'; '... die drei Personen, mit denen Sie
 näher befreundet sind und mit denen Sie sich am häufigsten treffen'.

Subjektive Schicht

22	Es wird heute viel über die verschiedenen Bevölkerungsschichten gesprochen. Welcher Schicht rechnen Sie sich selbst eher zu, der Unterschicht, der Arbeiterschicht, der Mittelschicht, der oberen Mittelschicht oder der Oberschicht?	Unterschicht 1 Arbeiterschicht 2 Mittelschicht 3 obere Mittelschicht 4 Oberschicht 5 <hr/> <i>INT.:</i> keiner dieser Schichten . . 6 <i>nicht</i> weiß nicht 8 <i>vorle-</i> Einstufung abgelehnt 7 <i>sen</i>	80
			9

Relative Deprivation

23	Im Vergleich dazu, wie Andere hier in der Bundesrepublik leben: glauben Sie, daß Sie Ihren gerechten Anteil an den angenehmen Dingen des Lebens erhalten, mehr als Ihren gerechten Anteil, etwas weniger oder sehr viel weniger?	gerechten Anteil 1 mehr als gerechten Anteil 2 etwas weniger 3 sehr viel weniger 4 <hr/> <i>INT.:</i> weiß nicht 8 <i>nicht</i> <i>vorle-</i> <i>sen</i>	10
			9

Klassenkonflikt (Split-Version B)

24	<u>INT.: grüne Liste 7 vorlegen</u>					
	Es wird oft gesagt, daß es Interessengegensätze zwischen verschiedenen Gruppen in der Bundesrepublik gibt, zum Beispiel zwischen politischen Gruppen, zwischen Männern und Frauen usw. Die Gegensätze sind aber nicht alle gleich stark. Ich will Ihnen nun einige solcher Gruppen nennen. Sagen Sie mir bitte, ob diese Gegensätze Ihrer Meinung nach sehr stark, ziemlich stark, eher schwach sind, oder ob es da gar keine Gegensätze gibt?					
	<u>INT.: vorlesen</u>					
	Gegensätze ...					
		sehr stark	ziemlich stark	eher schwach	gibt gar keine	weiß nicht
	- zwischen politisch links und politisch rechts stehenden Leuten	1	2	3	4	8
	- zwischen Arbeitgebern und Arbeitnehmern	1	2	3	4	8
	- zwischen Leuten mit Volksschulbildung und Akademikern	1	2	3	4	8
	- zwischen Leuten mit und Leuten ohne Kinder	1	2	3	4	8
	- zwischen Jungen und Alten	1	2	3	4	8
	- zwischen Arm und Reich	1	2	3	4	8
+)	- zwischen Erwerbstätigen und Rentnern	1	2	3	4	8
	- zwischen Politikern und den einfachen Bürgern	1	2	3	4	8
	- zwischen Kapitalisten und Arbeiterklasse	1	2	3	4	8
	- zwischen Gastarbeitern und Deutschen	1	2	3	4	8
	- zwischen Männern und Frauen	1	2	3	4	8

Version A: 'Interessenkonflikte' statt 'Interessengegensätze'

+) nur diese Items wurden in die Auswertung einbezogen

Bildung

S3	<u>AN ALLE:</u>			
	Welchen allgemeinbildenden Schulabschluß haben Sie? Einen Volks- oder Hauptschulabschluß, Mittlere Reife oder Realschulabschluß, die Fachhochschulreife, das Abitur oder keinen dieser Abschlüsse.			
	<u>INT.: nur den höchsten Abschluß angeben lassen</u>			
	Volks-/ Hauptschulabschluß	1	73	
	Mittlere Reife, Realschulabschluß (Fachschulreife)	2		
	Fachhochschulreife (Abschluß einer Fachoberschule etc.)	3		
	Abitur (Hochschulreife)	4		
	keinen dieser Abschlüsse	5		9

Privatisierung (= kollektive Philosophie bei Centers)

31 INT.: rosa Liste 10 vorlegen

Wir würden gerne noch Ihre Ansicht über einige wichtige politische Probleme hören.
Sagen Sie mir bitte mit Hilfe dieser Skala, in welchem Ausmaß Sie den folgenden Aussagen zustimmen. Die Zahl 1 bedeutet, daß Sie der Aussage überhaupt nicht zustimmen, die Zahl 7 drückt aus, daß Sie voll und ganz zustimmen. Mit den Zahlen dazwischen können Sie Ihre Antwort abstimmen.

		INT.: hier Skalenwert notieren	
A	Die Kernenergie sollte weiterentwickelt werden, um unseren zukünftigen Energiebedarf decken zu können	- - - - -	40/41
B	Für terroristische Gewalttaten sollte die Todesstrafe eingeführt werden	- - - - -	42/43
C	Mehr öffentliche Dienstleistungen, wie z.B. Müllabfuhr, Post oder Verkehrsbetriebe sollten von privaten Unternehmen übernommen werden	- - - - -	44/45
D	Frauen sollten über eine Abtreibung selbst entscheiden können	- - - - -	46/47

Parteipräferenz

34 INT.: blaue Liste 12 "Thermometer" (+5 bis -5) überreichen

Mich würde interessieren, wie Sie heute über die folgenden Parteien CDU, SPD, CSU, FDP, NPD, DKP und "Die Grünen" denken.
Hier haben wir eine Art Thermometer. Je weiter Sie auf den weißen Kästchen nach oben gehen, umso sympathischer ist Ihnen die betreffende Partei; je weiter Sie auf den schwarzen Kästchen nach unten gehen, umso mehr lehnen Sie die Partei ab. Fangen wir mit der CDU an.

INT.: Bei jeder Partei Thermometerstufe einkreisen;
auf "+" und "-" achten!

A	Ganz allgemein: was halten Sie von der CDU?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	52/53
B	Und was halten Sie von der SPD?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	54/55
C	Und was halten Sie von der CSU?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	56/57
D	Und was halten Sie von der FDP?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	58/59
E	Und was halten Sie von der NPD?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	60/61
F	Und was halten Sie von der DKP?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	62/63
G	Und was halten Sie von den "Grünen"?	+5 +4 +3 +2 +1 0 -1 -2 -3 -4 -5	64/65

+) nur diese Items wurden in die Auswertung einbezogen

Einkommen

S 39	Wir hätten nun gerne von Ihnen genauer gewußt, welche Personen hier in Ihrem Haushalt leben. Beginnen wir mit Ihnen. Nennen Sie mir bitte Ihren Vornamen.	
A	<p><u>INT.:</u> Vornamen des Befragten in das erste Kästchen unter A 1 eintragen</p> <p>Falls Einpersonenhaushalt, dann gleich weiter mit den Fragen 39F, 39G</p> <p>Nennen Sie mir bitte nun noch die Vornamen der anderen Haushaltsmitglieder, damit ich eine vollständige Liste der Personen habe, die hier leben. Beginnen Sie bitte mit der Ältesten Person und gehen Sie dem Alter nach weiter.</p> <p><u>INT.:</u> Alle Vornamen weiter unter A (2, 3, 4 usw.) eintragen. Nachfassen:</p> <p>Haben Sie auch niemanden vergessen? Zum Beispiel jemand, der normalerweise hier wohnt, aber zur Zeit abwesend ist, im Krankenhaus, Ferien usw. Haben Sie auch an die Kleinkinder gedacht, die zum Haushalt gehören?</p>	<p>→ Spalte A</p>
S 39	<p><u>INT.:</u> weiße Liste S5 vorlegen</p> <p>In welchem Verwandtschaftsverhältnis zu Ihnen (oder nicht verwandt) stehen die Personen, die Sie mir eben genannt haben? Bitte nennen Sie die entsprechende Ziffer.</p> <p><u>INT.:</u> Entsprechende Ziffer unter B für <u>jedes</u> Haushaltsmitglied eintragen</p>	<p>→ Spalte B</p>
S 39	<p><u>INT.:</u> Geschlecht kennzeichnen (unter C einkreisen). Nur bei solchen Personen nachfragen, bei denen aus dem Vornamen nicht ersichtlich ist, ob männlich oder weiblich.</p>	<p>→ Spalte C</p>
S 39	<p>Sagen Sie mir bitte jetzt, in welchem Jahr die betreffenden Personen geboren sind.</p> <p><u>INT.:</u> Unter D für <u>jedes</u> Haushaltsmitglied eintragen</p>	<p>→ Spalte D</p>
S 39	<p><u>INT.:</u> weiße Liste S6 vorlegen</p> <p>Können Sie mir bitte noch etwas über den Familienstand der Personen in Ihrem Haushalt sagen?</p> <p><u>INT.:</u> Unter E für <u>jedes</u> Haushaltsmitglied eintragen</p>	<p>→ Spalte E</p>
S 39	<p>Sagen Sie mir bitte noch für jede Person im Haushalt, ob sie über ein eigenes Einkommen verfügt? Beginnen wir wieder mit Ihnen.</p> <p><u>INT.:</u> Für jede Person einkreisen, ob Sie über eigene Einkünfte Einkünfte verfügt - Befragte(n) selbst nicht vergessen!</p>	<p>→ Spalte F</p>
S 39	<p><u>INT.:</u> Für jede Person mit eigenen Einkünften weiterfragen:</p> <p>Wie hoch ist das monatliche Netto-Einkommen von Ich meine die Summe, die nach Abzug der Steuern und Sozialversicherungsbeiträge übrigbleibt?</p> <p><u>INT.:</u> Höhe des Netto-Einkommens unter Spalte G eintragen - Wenn Einpersonenhaushalt, dann folgende Frage stellen:</p> <p>"Wie hoch ist Ihr monatliches Netto-Einkommen?"</p> <p>Ich meine die Summe, die nach Abzug der Steuern und Sozialversicherungsbeiträge übrigbleibt."</p>	<p>→ Spalte G</p>

+)

S39 A	S39 B	S39 C	S39 D	S39 E	S39 F	S39 G
VORNAME	INT.: Liste S5 Verwandtschaftsgrad zum Befragten	Geschlecht männ- weib- lich lich	Geburtsjahr	INT.: Liste S6 Familienstand verhel- getr. ver- gesch. ledig ratet lebend witw.	eigenes Einkommen	Höhe des eigenen monatlichen Netto-Einkommens
1	Befragter				ja . . 1 nein . 2	29 30/31/32/33 DM 9999
INT.: WENN NUR EINE PERSON IM HAUSHALT, WEITER MIT FRAGE → S 40						
2	34/75	36 1 2	37/38/39/40 9999	41 1 2 3 4 5	42 ja . . 1 nein . 2	43/44/45/46 DM 9999
3	47/48	49 1 2	50/51/52/53 9999	54 1 2 3 4 5	55 ja . . 1 nein . 2	56/57/58/59 DM 9999
4	60/61	62 1 2	63/64/65/66 9999	67 1 2 3 4 5	68 ja . . 1 nein . 2	69/70/71/72 DM 9999
5	73/74	75 1 2	76/77/78/79 9999	80 1 2 3 4 5	81 ja . . 1 nein . 2	82/83/84/85 DM 9999

•
•
•

+) nur dieses Item wurde in die Auswertung einbezogen

Links - Rechts

32	<p>Viele Leute verwenden die Begriffe LINKS und RECHTS, wenn es darum geht, unterschiedliche politische Einstellungen zu kennzeichnen. Wir haben hier einen Maßstab, der von links nach rechts verläuft. Wenn Sie an Ihre eigenen politischen Ansichten denken, wo würden Sie diese Ansichten auf dieser Skala einstufen? Machen Sie bitte ein Kreuz in eines der Kästchen.</p> <p><u>INT.:</u> Stift überreichen. Befragter soll selbst auf der nächsten Seite ankreuzen. <u>Unbedingt</u> darauf achten, daß <u>innerhalb</u> eines Kästchens angekreuzt wird.</p>
----	--

LINKS	<div style="position: absolute; top: 0; right: 0; bottom: 0; left: 0;"></div>	RECHTS
-------	---	--------

Oben - Unten

346	<p>Ich komme nun zur letzten Frage</p> <p>In unserer Gesellschaft gibt es Bevölkerungsgruppen, die eher oben stehen und solche, die eher unten stehen. Wir haben hier eine Skala, die von oben nach unten verläuft. Wenn Sie an sich selbst denken: wo auf dieser Skala würden Sie sich einordnen?</p> <p><u>INT.:</u> Bleistift überreichen, Befragter soll selbst auf der nächsten Seite Skala ankreuzen. <u>Unbedingt</u> darauf achten, daß <u>innerhalb</u> eines Kästchens angekreuzt wird</p>
-----	--

OBEN	<div style="position: absolute; top: 0; right: 0; bottom: 0; left: 0;"></div>	UNTEN
------	---	-------

Berufliche Stellung

S6	<p><u>INT.:</u> weiße Liste S3 vorlegen</p> <p>Bitte ordnen Sie Ihre berufliche Stellung nach dieser Liste ein.</p>	<p>bei Kennziffer ^{77/78} 10 bis 23 → S7a</p> <p>bei Ziffer 30 → S7b</p> <p>bei Kennziffer 40 bis 74 → S7c</p>
	<p>-----</p> <p>Kennziffer eintragen</p>	

LISTE S3

U 805/79

zu Frage S6, S16, S29, S33

<u>SELBSTÄNDIGE LANDWIRTE</u> mit landwirtschaftlich genutzter Fläche von ...	<u>AKADEMISCHE FREIE BERUFE</u> (z.B. Arzt, Rechtsanwalt mit eigener Praxis)	<u>SONSTIGE SELBSTÄNDIGE</u> außerhalb der Landwirtschaft mit ...
10 - unter 10 ha 11 - 10 ha bis unter 20 ha 12 - 20 ha und mehr	15 - 1 Mitarbeiter oder allein 16 - 2 - 9 Mitarbeiter 17 - 10 Mitarbeiter und mehr	21 - 1 Mitarbeiter oder allein 22 - 2 - 9 Mitarbeiter 23 - 10 Mitarbeiter und mehr
30 <u>MITHELFFENDE FAMILIENANGEHÖRIGE</u>		
<u>BEAMTE/ RICHTER/ BERUFSSOLDATEN</u> 40 Beamte im einfachen Dienst (bis einschl. Oberamtsmeister) 41 Beamte im mittleren Dienst (vom Assistenten bis einschl. Hauptsekretär/ Amtsinspektor) 42 Beamte im gehobenen Dienst (vom Inspektor bis einschl. Oberamtmann/ Oberamtsrat) 43 Beamte im höheren Dienst, Richter (vom Regierungsrat aufwärts)	<u>ANGESTELLTE</u> 50 Industrie- und Werkmeister im Angestelltenverhältnis 51 Angestellte mit einfacherer Tätigkeit (z.B. Verkäufer, Kontorist, Stenotypistin) 52 Angestellte, die schwierigere Aufgaben nach allgemeiner Anweisung selbständig erledigen (z.B. Sachbearbeiter, Buchhalter, technischer Zeichner) 53 Angestellte, die selbständige Leistungen in verantwortungs- voller Tätigkeit erbringen oder begrenzte Verantwortung für die Tätigkeit anderer tragen (z.B. wiss. Mitarbeiter, Prokurist, Abteilungsleiter) 54 Angestellte mit umfassenden Führungsaufgaben und Entscheidungsbefugnissen (z.B. Direktor, Geschäftsführer, Vorstand größerer Betriebe und Verbände)	
49 <u>WEHRPFLICHTIGE</u>		
<u>ARBEITER</u> 60 ungelernte Arbeiter 61 angelernte Arbeiter 62 gelernte und Facharbeiter 63 Vorarbeiter und Kolonnenführer 64 Meister/ Poliere	<u>IN AUSBILDUNG</u> 70 kaufmännische/ Verwaltungs-Lehrlinge 71 gewerbliche Lehrlinge 72 haus-/ landwirtschaftliche Lehrlinge 73 Beamtenanwärter/ Beamte im Vorbereitungsdienst 74 Praktikanten/ Volontäre	

Anhang 2: I. PEARSON Korrelationsmatrix

	y ₁	y ₂	y ₃	y ₄	y ₅	y ₆	y ₇	y ₈	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆
y ₁	1.000													
y ₂	.370	1.000												
y ₃	.030	.090	1.000											
y ₄	-.070	-.050	-.060	1.000										
y ₅	-.060	-.070	-.020	.390	1.000									
y ₆	-.060	-.050	-.030	.410	.510	1.000								
y ₇	.020	.030	.050	-.060	-.040	-.070	1.000							
y ₈	.240	.260	.070	-.120	-.110	-.120	-.010	1.000						
x ₁	.300	.270	.020	-.040	-.050	-.040	.100	.150	1.000					
x ₂	.420	.180	-.110	-.010	-.060	-.070	.010	.140	.280	1.000				
x ₃	.480	.260	.020	-.050	-.090	-.050	.090	.200	.530	.520	1.000			
x ₄	.330	.220	.010	-.050	-.040	-.020	.040	.130	.280	.360	.450	1.000		
x ₅	.320	.180	-.020	.050	.010	.010	.020	.100	.270	.380	.420	.410	1.000	
x ₆	.340	.170	-.040	-.020	-.020	-.030	.010	.100	.320	.400	.490	.430	.410	1.000

EINKOMMENSDISKRIMINIERUNG VON FRAUEN¹

MESSUNG, ANALYSEVERFAHREN UND EMPIRISCHE ANWENDUNGEN AUF ANGESTELLTENEINKOMMEN IN DER BUNDESREPUBLIK

Andreas Diekmann

1. Einleitung

Trotz aller formalen Gleichberechtigungspostulate sind in westlichen als auch östlichen Industriegesellschaften erhebliche Verdienstspannen zwischen Frauen- und Männer-Einkommen nachweisbar (z.B. Schmidt 1978; Pettmann 1975; Kyn 1978; Swafford 1978). In der Bundesrepublik und in Österreich liegen die Einkommen der Männer im Durchschnitt ungefähr 50 Prozent höher als die der Frauen. Dieser prozentuale Abstand hat sich hartnäckig und langfristig in verschiedensten Bereichen gehalten, so daß zwar vereinfachend, aber treffend von einer "50-Prozent-Regel" (Christl und Wagner 1982) gesprochen werden kann.

Wie sind diese Unterschiede zu erklären? Angehörige einer diskriminierten Bevölkerungsgruppe sind in den meisten Fällen in doppelter Weise benachteiligt. Da sich Diskriminierungen in frühen Phasen des Lebenszyklus kumulieren (Duncan 1968) und auf spätere Stadien auswirken, bringen Angehörige diskriminierter Gruppen ungünstigere Voraussetzungen auf dem Arbeitsmarkt mit. Darüber hinaus sind sie aber häufig noch zusätzlich benachteiligt, auch wenn sie die gleichen Voraussetzungen aufweisen wie Mitglieder der privilegierten Gruppe. Diese beiden Hauptkomponenten der Einkommensunterschiede wollen wir in Abschnitt 1 etwas genauer betrachten. In Teil 2 wird ein Verfahren zur Messung der Diskriminierungskomponenten entwickelt, und in Teil 3 wird das Verfahren zur Analyse von amtlichen Daten aus dem Angestelltenbereich im Zeitraum 1970 bis 1980 angewandt. Eine multivariate Erweiterung der Analyse auf der Basis von Umfragedaten wird in Teil 4 diskutiert.

2. Arten der Diskriminierung

Wenn allgemein gesprochen das Innehaben einer Position mit Belohnungen verbunden ist und weiterhin zwischen einer privilegierten und einer unterprivilegierten Personengruppe im Mittel eine Belohnungsdifferenz besteht, dann kann dieser Unterschied das Resultat zweier verschiedener Mechanismen der Diskriminierung sein. Zum einen können Mitglieder der privilegierten Gruppe bessere Zugangschancen zu einer begehrten Position haben, zum anderen kann es der Fall sein, daß Mitglieder der privilegierten Gruppe bei gleicher Position höhere Belohnungen erhalten. Die Gruppenzugehörigkeit wird dabei nicht durch die Belohnungshöhe festgelegt, sondern durch ein unabhängig hiervon erkennbares Merkmal wie Rasse, Geschlecht, Nationalität, Behindertenstatus usw. Positionen und Belohnungen können sehr allgemein verstanden werden. Zum Beispiel mag die "Position" ein Restaurant sein und die "Belohnung" der Service. Werden Menschen mit bestimmter Hautfarbe an der Tür abgewiesen, dann handelt es sich um Diskriminierung der ersten Art (externe Diskriminierung). Werden sie hineingelassen, aber schlechter bedient, dann ist es Diskriminierung der zweiten Art (interne Diskriminierung).

Bei der Analyse der Einkommensdiskriminierung von Frauen kann die "Position" eine Branche, eine Berufsgruppe, ein Beruf oder eine spezielle Tätigkeit sein. Die Belohnung ist in diesem Fall das mit der Position verbundene Einkommen. Externe oder Zugangschancendiskriminierung existiert, wenn die Verteilung der Frauen auf die Positionen sich von der Verteilung der Männer derart unterscheidet, daß die Chance, einträgliche Positionen einzunehmen, geringer ist, während es sich bei den wenig einträglichen Positionen genau umgekehrt verhält. Interne Diskriminierung liegt dagegen vor, wenn Frauen in der gleichen Position wie Männer weniger Einkünfte beziehen.

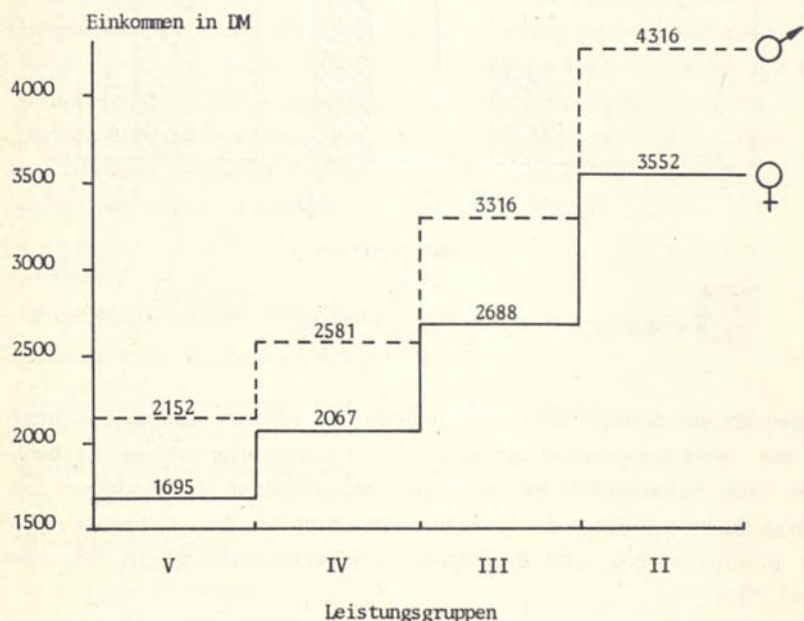
Wir führen diese neuen Begriffe ein, um ein allgemeines Diskriminierungsmodell zu formulieren. Wie weiter unten ausgeführt, können sich Positionen auch auf mehrdimensionale Eigenschaften beziehen, z.B. auf Bildungs-Berufs-Positionen. Wenn jedoch im einfacheren Fall Positionen Berufe oder bestimmte Tätigkeiten darstellen, entspricht in einer vielleicht vertrauteren Terminologie die Zugangschancendiskriminierung der Beschäftigungsdiskriminierung ("job discrimination") und die interne Diskriminierung der "direkten" Einkommensdiskriminierung.

Im folgenden Abschnitt wird ein Verfahren angegeben, welches gestattet, die quantitative Höhe der Zugangschancendiskriminierung und der internen Diskriminierung zu bestimmen. Als Datenbasis ist dafür die Kenntnis der Positionsverteilungen und der mit den Positionen verbundenen Einkünfte jeweils für beide Geschlechter erforderlich.

3. Ein Verfahren zur Messung der Diskriminierung

Die Vorgehensweise sei an amtlichen Einkommensdaten für Angestellte in der Industrie aus dem Jahre 1980 demonstriert. In Schaubild 1 sind die Einkommensunterschiede für vier Leistungsgruppen (II = höchste Leistungsgruppe, V = unterste Leistungsgruppe)² dargestellt.

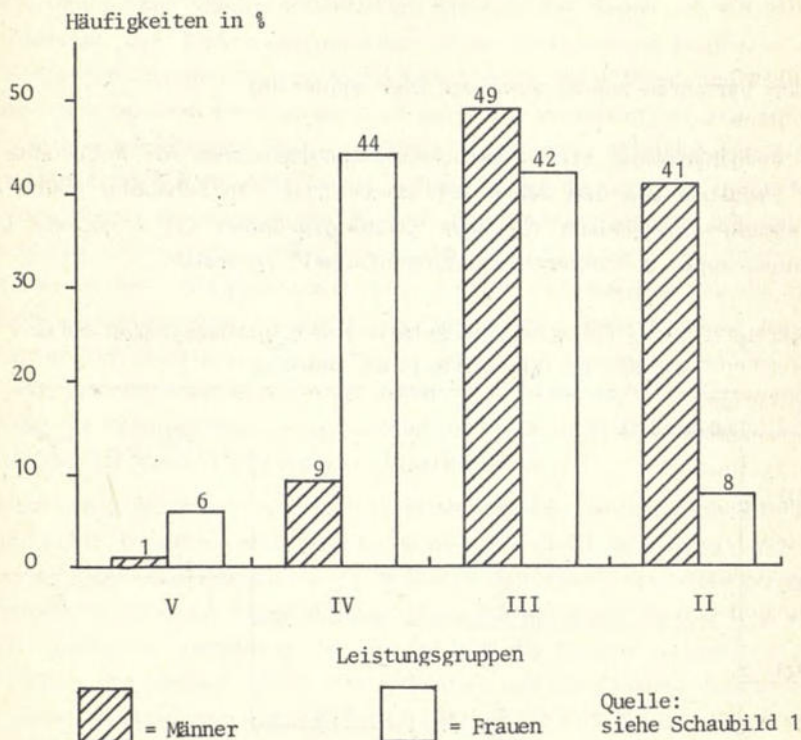
Schaubild 1: Bruttoverdienste (Tarifgehälter) nach Leistungsgruppen und Geschlecht für Angestellte in der Industrie



Quelle: Amtliche Verdienststatistik des Statistischen Bundesamtes Wiesbaden, Fachserie 16, Reihe 4.2

Im vorliegenden Fall sind die "Positionen" die Leistungsgruppen. Die geschlechtsspezifischen Verteilungen zeigt Schaubild 2.

Schaubild 2: Verteilung auf Leistungsgruppen nach Geschlecht



Aus den beiden Schaubildern geht anschaulich hervor, daß sowohl innerhalb der Leistungsgruppen erhebliche Einkommensunterschiede zu beobachten sind (Schaubild 1) als auch die Tatsache einer ungünstigeren Leistungsgruppenverteilung der Frauen (Schaubild 2). Es existiert also interne Diskriminierung und Zugangschancendiskriminierung (im folgenden ID und ZD).

Wie lassen sich nun die Anteile der beiden Diskriminierungstypen quantitativ bestimmen? Dazu ist es zunächst erforderlich, zwei genauere Definitionen anzugeben.

Definition 1: Interne Diskriminierung

Der Betrag von ID entspricht dem hypothetischen Pro-Kopf-Mehrgewinn der diskriminierten Gruppe bei gegebener Positionsverteilung, wenn Angehörige der diskriminierten Gruppe in jeder Position die gleichen Einkünfte wie Angehörige der privilegierten Gruppe beziehen.

Definition 2: Zugangschancendiskriminierung

Die Höhe von ZD bestimmt sich als hypothetischer Mehrgewinn pro Kopf der diskriminierten Gruppe, der dadurch erzielt wird, daß die Positionsverteilung der diskriminierten Gruppe, an die Verteilung der privilegierten Gruppe angepaßt wird bei gegebenen internen Einkommensunterschieden.

Anders ausgedrückt sind ID und ZD Antworten auf die folgenden hypothetischen "Was-wäre-wenn-Fragen": Wie ändert sich das Einkommen eines Angehörigen der diskriminierten Gruppe im Durchschnitt, wenn keine Einkommensunterschiede innerhalb einer Leistungsklasse existierten (ID-Effekt)? Und wie ändert sich der Einkommens-Erwartungswert, wenn Geschlechtsunterschiede der Zugangschancen beseitigt werden bei unveränderten Entlohnungsregeln pro Leistungsklasse (ZD-Effekt)? Oder: ID ist der Verlust aufgrund "reiner" interner Diskriminierung, ZD der Verlust aufgrund "reiner" Zugangschancendiskriminierung.

Notation

Führen wir folgende Bezeichnungen ein:

Durchschnitts-Einkommen Frauen bzw. Männer: \bar{Y}^w, \bar{Y}^m

Durchschnitts-Einkommen Frauen bzw. Männer pro Leistungsklasse i \bar{Y}_i^w, \bar{Y}_i^m

Relative Häufigkeiten Position i für Frauen bzw. Männer f_i^w, f_i^m

Zahl der Positionen n

Anzahl der Personen: N

Zahl der Frauen bzw. Männer: N^w, N^m

Hypothetisches Pro-Kopf-Einkommen der Frauen, wenn kein Lohnunterschiede innerhalb der Leistungsklassen bestehen (Mittelwert): \bar{Y}_{H1}^w

Hypothetisches Pro-Kopf-Einkommen der Frauen,
wenn gleiche Positionsverteilung wie Männer
(Mittelwert):

$$\bar{Y}_{H2}^w$$

Dann gilt gemäß Definition 1:

$$(1) \quad ID = \bar{Y}_{H1}^w - \bar{Y}^w$$

ID ist also der hypothetische Mehrgewinn, den Frauen über ihr tatsächliches Einkommen hinaus erzielen, wenn sie innerhalb jeder Leistungsklasse im Mittel in gleicher Weise wie die Männer entlohnt werden würden.

Nach Definition 2 ergibt sich:

$$(2) \quad ZD = \bar{Y}_{H2}^w - \bar{Y}^w$$

ZD entspricht demnach dem hypothetischen Mehrgewinn, den Frauen erhielten, wenn die Aufteilung auf die Leistungsgruppen der männlichen Verteilung entspräche.

Nun ist $\bar{Y}_{H1}^w = \sum f_i^w \bar{y}_i^m$, wobei die Summierung hier und im folgenden über alle Positionen $i = 1$ bis n erfolgt. Daher kann für ID geschrieben werden:

$$(3) \quad ID = \sum f_i^w \bar{y}_i^m - \bar{Y}^w$$

Und entsprechend gilt für ZD, wenn $\bar{Y}_{H2}^w = \sum f_i^m \bar{y}_i^w$ berücksichtigt wird:

$$(4) \quad ZD = \sum f_i^m \bar{y}_i^w - \bar{Y}^w$$

Das Durchschnittseinkommen der Frauen \bar{Y}^w ist mit dem Ausdruck $\sum f_i^w \bar{y}_i^w$ identisch. Daher können (3) und (4) auch in folgender Weise geschrieben werden:

$$(5) \quad ID = \sum f_i^w (\bar{y}_i^m - \bar{y}_i^w)$$

$$(6) \quad ZD = \sum \bar{y}_i^w (f_i^m - f_i^w)$$

ID und ZD lassen sich also auch als gewichtete Differenzen interpretieren. Im ersten Fall werden die Einkommensdifferenzen jeder Position mit der Positionshäufigkeit gewichtet und aufsummiert, im zweiten Fall werden die Chancendifferenzen (Differenzen der Positionshäufigkeiten) mit dem Einkommen gewichtet und summiert.

Der globale Einkommensunterschied ($\bar{Y}^m - \bar{Y}^w$) ist normalerweise größer als die Summe aus ID und ZD.³ Dies ist dadurch zu erklären, daß neben den "reinen" Effekten von ID und ZD noch ein simultaner Effekt existiert. Dieser Rest- oder Interaktionseffekt gibt an, welchen Betrag die diskriminierte Gruppe durchschnittlich über die "reinen" Effekte hinaus erhalten würde, wenn gleichzeitig die Frauen-Einkommen je Position dem männlichen Einkommen und die Zugangschancen der Frauen den männlichen Zugangschancen angepaßt werden. Verwenden wir für den Interaktionseffekt das Kürzel "IE", dann läßt sich folgende Zerlegungsgleichung formulieren:

$$(7) \quad (\bar{Y}^m - \bar{Y}^w) = ID + ZD + IE$$

Die Größen auf der rechten Seite lassen sich auch in Prozentanteilen am globalen Einkommensunterschied ausdrücken. Die prozentualen Anteile lauten somit:

$$(8) \quad ID \% = \frac{ID * 100}{(\bar{Y}^m - \bar{Y}^w)}$$

$$(9) \quad ZD \% = \frac{ZD * 100}{(\bar{Y}^m - \bar{Y}^w)}$$

$$(10) \quad IE \% = \frac{IE * 100}{(\bar{Y}^m - \bar{Y}^w)}$$

Die drei Prozentwerte müssen sich natürlich zu 100 Prozent addieren. Multipliziert man ferner die Zerlegungsgleichung (7) mit der absoluten Zahl der Frauen N^w , dann erhält man eine Zerlegungsgleichung für die gesamte Einkommenssumme, die die diskriminierte Population aufgrund interner und externer Diskriminierung als Verlust verbuchen muß. Die Gleichung lautet wie folgt:

$$(11) \quad N^W (\bar{Y}^m - \bar{Y}^W) = N^W * ID + N^W * ZD + N^W * IE$$

Mögliche Diskriminierungskombinationen

In vielen Fällen sind - wie bei der Einkommensdiskriminierung von Frauen - diskriminierte Gruppen in doppelter Weise benachteiligt, d.h. es existiert interne und externe Diskriminierung. Diese Kombination muß allerdings nicht immer zutreffen. So mag es vorkommen, daß eine diskriminierte Gruppe im Durchschnitt zwar weniger verdient, aber dennoch privilegiert ist in Hinblick auf die Zugangschancen. In einem derartigen Fall wäre ZD negativ und ID könnte Werte annehmen, die höher als der globale Einkommensunterschied liegen. Die interne Diskriminierung übersteigt dann die Vorteile bezüglich der Zugangschancen. Auch die umgekehrte Kombination ist denkbar: ID ist negativ und ZD ist positiv, wobei ein höherer Wert externer Diskriminierung als interner Privilegierung unterstellt wird, da ansonsten kein Rückstand mehr bezüglich des globalen Durchschnittseinkommens der diskriminierten Gruppe feststellbar wäre.

Zum Diskriminierungsbegriff

Bisher wurde der gesamte Einkommensabstand $\bar{Y}^m - \bar{Y}^W$ mit dem Etikett "Diskriminierung" versehen. Dagegen könnte das Argument eingewandt werden, daß die globale Einkommensdifferenz nur dann auf Diskriminierung beruhte, wenn folgende Umstände für beide Geschlechter in jeder Leistungsklasse gegeben sind: Gleiche Arbeitszeit, gleiche Schwere der Arbeit, gleiche Qualifikation, gleiche Berufserfahrung etc. - oder nach einer üblichen Definition der ökonomischen Diskriminierungsliteratur - gleiche Produktivität. Wir wollen die Berechtigung dieses Einwands nicht abstreiten, nur handelt es sich dabei zunächst um einen Streit über Namen (der allerdings nicht unwichtig ist, da damit meist auch Wertungen verbunden sind). Die erläuterte Strategie der Zerlegung des globalen Einkommensunterschieds ist natürlich auch dann gültig und sinnvoll, wenn man sich dagegen sträubt, die Komponenten als "Diskriminierung" zu bezeichnen. Darüber hinaus können aber auch die erwähnten weiteren Variablen von dem Modell berücksichtigt werden, wie im folgenden zu zeigen ist. Auf diese Weise wird die Möglichkeit geschaffen, je nach Diskriminierungsdefinition die relevanten Variablen einzubeziehen und die Diskriminierungskomponente zu isolieren.

Verallgemeinerung auf den multivariaten Fall

In unserem Beispiel waren die Positionen Ausprägungen der einkommensbestimmenden Variable "Leistungsgruppen". Die Einbeziehung weiterer Variablen wie z.B. der Bildungsstufe kann durch den einfachen Trick der Konstruktion "komplexer" Positionen bewerkstelligt werden. D.h. im multivariaten Fall sind die Positionen Kombinationen von Ausprägungen der berücksichtigten Variablen, z.B. Bildungs-Leistungsgruppen-Kombinationen. Die Zugangschance zu einer "komplexen Position", z.B. zu der Kombination Abitur - Leistungsgruppe III, entspricht dann dem entsprechenden Wert der gemeinsamen Häufigkeitsverteilung von Bildung und Leistungsgruppe. Sofern es die Datenlage gestattet, kann mit dem angegebenen Verfahren die Analyse beliebig verfeinert werden.

Die Verallgemeinerung der Formel für ID lautet dann:

$$(12) \quad ID = \sum_{i_1=1}^{n_1} \dots \sum_{i_r=1}^{n_r} f^w(i_1, \dots, i_r) [\bar{y}^m(i_1, \dots, i_r) - \bar{y}^w(i_1, \dots, i_r)]$$

Hierbei sind n_1, \dots, n_r die Zahlen der Kategorien der jeweils berücksichtigten Variablen, $f^w(i_1, \dots, i_r)$ ist die relative Häufigkeit in der entsprechenden Ausprägung der gemeinsamen Häufigkeitsverteilung der Frauen und $\bar{y}^m(i_1, \dots, i_r)$ bzw. $\bar{y}^w(i_1, \dots, i_r)$ sind die Einkommensmittelwerte der Frauen bzw. Männer für die in Frage stehende Merkmalskombination. Formel (12) ist die Verallgemeinerung von (5). In völlig analoger Weise läßt sich auch Gleichung (6) erweitern:

$$(13) \quad ZD = \sum_{i_1=1}^{n_1} \dots \sum_{i_r=1}^{n_r} \bar{y}^w(i_1, \dots, i_r) [f^m(i_1, \dots, i_r) - f^w(i_1, \dots, i_r)]$$

Werden die Variablen danach ausgewählt, daß die Merkmalskombinationen produktivitäts-homogen sind, d.h. daß alle Inhaber der Position unabhängig vom Geschlecht gleiche Produktivität aufweisen, dann stellt (12) eine Diskriminierungsmessung auf der Basis der ökonomischen Produktivitäts-Definition dar.

Als Schätzungen der Produktivität werden häufig die Variablen Ausbildung, Berufserfahrung usw. herangezogen. Wenn auch bei Kontrolle die-

ser Variablen ID positive Werte aufweist, dann kann wohl mit größerer Berechtigung von direkter Diskriminierung gesprochen werden. Darüber hinaus verdient aber auch die Analyse der indirekten Diskriminierung nach dem hier vertretenen Standpunkt besondere Aufmerksamkeit.

Regressionsanalyse als Spezialfall

Bei den bisherigen Ausführungen wurde keine Einschränkung bezüglich des Skalenniveaus der einkommensbestimmenden Variablen gemacht. Erforderlich ist nur, daß das Einkommen - bzw. allgemein gesprochen die Belohnung für die Positionsinhaber - quantitativ meßbar ist. Die erklärenden Variablen können dagegen qualitativen Charakter annehmen wie in unserem Fall die Variable "Leistungsgruppen".

Weisen jedoch auch die erklärenden Variablen mindestens Intervallskalenniveau auf, dann können unter der Annahme bestimmter Einkommensfunktionen Regressionsanalysen - getrennt für beide Geschlechter - durchgeführt werden⁴.

a) Das Verfahren von Duncan

Die Analyse von Diskriminierungsprozessen und die Schätzung eines Diskriminierungsanteils mittels Regressionsanalyse wurde insbesondere von Duncan (1968) propagiert⁵. Der Einfachheit halber illustrieren wir die folgenden Überlegungen für den bivariaten Fall. Die Resultate sind jedoch auf multivariate Analysen ohne weiteres übertragbar. Wenn wir von einer unabhängigen Variablen (x) - z.B. der Schulbildung gemessen in Jahren - und dem Einkommen (y) ausgehen, dann kann für die männliche und die weibliche Stichprobe je eine Regressionsgleichung geschrieben werden (\hat{y} ist die geschätzte Einkommenshöhe):

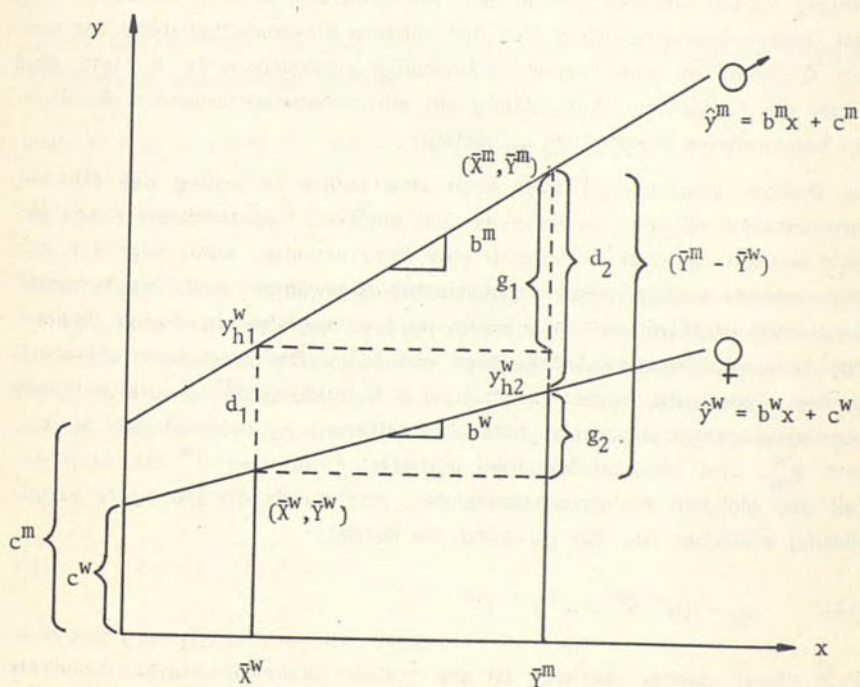
$$(14) \quad \hat{y}^m = b^m x + c^m$$

$$(15) \quad \hat{y}^w = b^w x + c^w$$

Im allgemeinen ist die privilegierte Gruppe in doppelter Weise begünstigt. Zum einen können strukturelle Privilegien existieren. Damit bezeichnen

wir die Situation, bei der erstens die männliche Einkommensfunktion oberhalb der weiblichen Funktion verläuft und zweitens auch die Steigung für die männliche Funktion größer ist als diejenige der weiblichen Funktion ($b^m > b^w$). In unserem Beispiel erzielen Bildungsinvestitionen von Männern dann einen höheren Einkommenszuwachs als Investitionen von Frauen. Zum anderen kann die privilegierte Gruppe bezüglich der einkommensbestimmenden Faktoren im Vorteil sein. Dies ist der Fall, wenn bei dem Beispiel die durchschnittliche Schulbildung der Männer (\bar{x}^m) größer ist als der mittlere Bildungsgrad der Frauen (\bar{x}^w). Schaubild 3 veranschaulicht die skizzierte Situation.

Schaubild 3: Graph der Regressionsgleichungen für die männliche und weibliche Stichprobe



Duncan (1967, 1968) versucht nun, eine "reine" Diskriminierungskomponente zu isolieren, indem er folgende hypothetische Frage stellt:

Angenommen, eine Frau mit durchschnittlicher Bildung \bar{X}^w erhielte ein Einkommen nach dem männlichen Entlohnungsmuster, so würde die Höhe des Einkommens in Schaubild 3 \hat{y}_{h1}^w Einheiten betragen. Den Wert erhält man, indem \bar{X}^w in die männliche Regressionsgerade (14) eingesetzt wird. Die Diskriminierungskomponente entspricht dann der Differenz zwischen dem hypothetischen Wert \hat{y}_{h1}^w in Schaubild 3 und dem tatsächlichen Einkommen \bar{Y}^w , also der Strecke d_1 . Formal ausgedrückt:

$$(16) \quad d_1 = (b^m \bar{X}^w + c^m) - \bar{Y}^w$$

Die Differenz zwischen dem globalen Einkommensunterschied ($\bar{Y}^m - \bar{Y}^w$) und d_1 ist die Restgröße g_1 . Diese Komponente des Einkommensabstands ist nach Duncan auf die Benachteiligung bezüglich der geringeren Schulbildung zurückzuführen, wenn wir bei unserem Beispiel bleiben. Nach dem Duncan-Verfahren wird also der mittlere Einkommensabstand der beiden Gruppen in eine "reine" Diskriminierungskomponente d_1 und eine durch die schlechtere Ausstattung mit einkommensbestimmenden Ressourcen begründbare Komponente g_1 zerlegt.

Wie Duncan bemerkt, ist aber auch eine zweite Zerlegung des Einkommensabstands möglich. Es kann nämlich auch die hypothetische Frage gestellt werden, welches Einkommen eine Frau erhielte, wenn zwar das diskriminierende weibliche Belohnungsmuster angewendet wird, die Schulbildung aber im Mittel dem männlichen Wert entspräche. In diesem Fall ergibt sich ein Einkommensschätzwert von \hat{y}_{h2}^w . Der Wert kann errechnet werden, indem die mittlere männliche Schulbildung \bar{X}^m in die weibliche Regressionsgerade eingesetzt wird. Die Differenz g_2 zwischen dem Schätzwert \hat{y}_{h2}^w und dem tatsächlichen mittleren Einkommen \bar{Y}^w ist dann der Teil des globalen Einkommensabstandes, der durch die geringere Schulbildung erklärbar ist. Für g_2 lautet die Formel:

$$(17) \quad g_2 = (b^w \bar{X}^m + c^w) - \bar{Y}^w$$

Nach dieser zweiten Methode ist die "reine" Diskriminierungskomponente d_2 das Komplement zu g_2 bezüglich des globalen Abstands ($\bar{Y}^m - \bar{Y}^w$).

Nach dem Duncan-Verfahren stehen also im Prinzip zwei Berechnungsmethoden zur Isolierung der "reinen" Diskriminierungskomponente zur Verfügung. Bei der ersten Methode werden die "weiblichen" Mittelwerte der unabhängigen Variablen in die männliche Regressionsgleichung, bei der zweiten Methode dagegen die "männlichen" Mittelwerte in die weibliche Regressionsgerade eingesetzt. Je nach der angewandten Methode erhält man in der Regel unterschiedliche Schätzungen der Diskriminierungskomponente, wobei d_2 im allgemeinen einen höheren Wert aufweist als d_1 .⁶ In dieser Situation ist die Kenntnis eines eindeutigen Kriteriums wünschenswert, das darüber Auskunft gibt, welche Diskriminierungsschätzung vorzuziehen ist. Wir werden im folgenden zeigen, daß unser allgemeineres Modell einen Weg weist, um diese Schwierigkeiten zu überwinden.⁷

b) Duncans Verfahren als Spezialfall des allgemeinen Diskriminierungsmodells - eine neue Deutung

Gemäß den Formeln (3) und (4) konnten für ID und ZD die folgenden Gleichungen geschrieben werden:

$$(18) \quad ID = \sum f_i^w \bar{y}_i^m - \bar{y}^w$$

$$(19) \quad ZD = \sum f_i^m \bar{y}_i^w - \bar{y}^w$$

Wenn nun die männlichen und weiblichen Entlohnungsmuster durch die jeweiligen Regressionsgeraden gegeben sind, dann können in (18) und (19) statt der Einkommensmittelwerte je Kategorie der unabhängigen Variable die Regressionsschätzwerte \hat{y}_i^m bzw. \hat{y}_i^w eingesetzt werden. Die resultierenden Gleichungen lauten also im Spezialfall der Regressionsanalyse:

$$(20) \quad ID = \sum f_i^w \hat{y}_i^m - \bar{y}^w$$

$$(21) \quad ZD = \sum f_i^m \hat{y}_i^w - \bar{y}^w$$

Nun gilt aber gemäß (14) und (15):

$$(22) \quad \hat{y}^m = b^m x + c^m$$

$$(23) \quad \hat{y}^w = b^w x + c^w$$

Wird in die rechte Seite der beiden Gleichungen in (20) und (21) \hat{y}_i^m und \hat{y}_i^w eingesetzt, so ergibt sich:

$$(24) \quad ID = b^m \sum f_i^w x_i + c^m \sum f_i^w - \bar{y}^w$$

$$(25) \quad ZD = b^w \sum f_i^m x_i + c^w \sum f_i^m - \bar{y}^w$$

Berücksichtigt man, daß $\sum f_i^w x_i = \bar{x}^w$, $\sum f_i^m x_i = \bar{x}^m$ und $\sum f_i^w = \sum f_i^m = 1$ ist, so vereinfachen sich (24) und (25) wie folgt:

$$(26) \quad ID = (b^m \bar{x}^w + c^m) - \bar{y}^w$$

$$(27) \quad ZD = (b^w \bar{x}^m + c^w) - \bar{y}^w$$

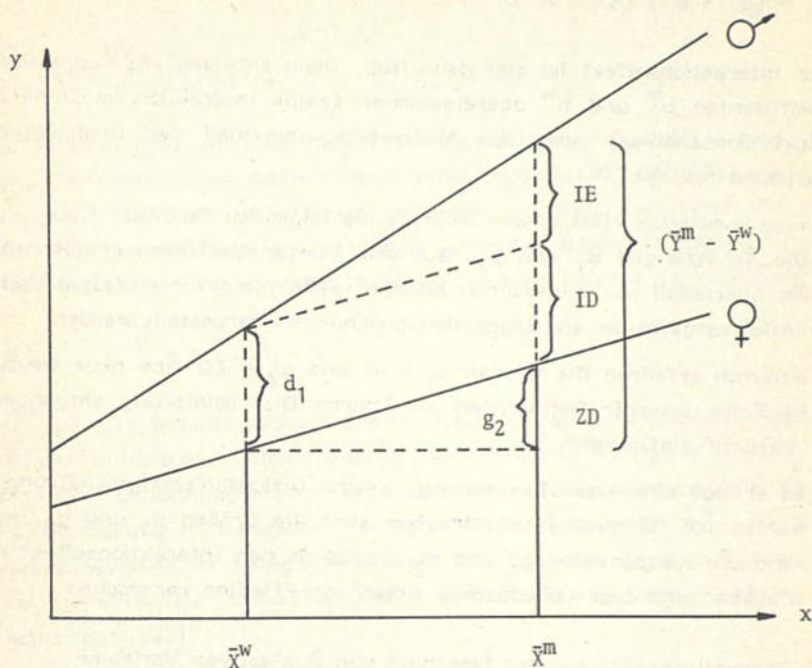
Ein Vergleich mit den Schätzungen nach dem Duncan-Verfahren (Formeln (16) und (17)) zeigt, daß ID mit d_1 und ZD mit g_2 identisch ist. Diese Überlegungen gelten auch bei mehreren unabhängigen Variablen, also im Falle multivariater Analyse.

Gemäß der Zerlegungsgleichung (7) gilt ferner:

$$(28) \quad (\bar{y}^m - \bar{y}^w) = ID + ZD + IE = d_1 + g_2 + IE$$

In Schaubild 4 wird die Zerlegung graphisch veranschaulicht:

Schaubild 4: Komponenten des Einkommensabstandes beim Regressionsmodell



Nach dem Duncan-Verfahren addieren sich d_1 und g_1 sowie d_2 und g_2 zu $(\bar{y}^m - \bar{y}^W)$. Unter Berücksichtigung von (28) erhält man daher die Beziehungen:

$$d_2 = d_1 + IE = ID + IE$$

$$g_1 = g_2 + IE = ZD + IE$$

Diese beiden Schätzungen schließen also jeweils den Interaktionseffekt ein und konfundieren daher zwei verschiedene Arten von Effekten. Normalerweise stellt d_2 eine Überschätzung der "reinen" Diskriminierung und g_1 eine Überschätzung der Zugangschancendiskriminierung dar.

Durch Umformungen (die mit elementarer Geometrie anhand der Schaubilder 3 und 4 nachvollziehbar sind) erhält man weiterhin die folgenden Ausdrücke:

$$ZD = b^w(\bar{X}^m - \bar{X}^w)$$

$$IE = (b^m - b^w)(\bar{X}^m - \bar{X}^w)$$

Der Interaktionseffekt ist also dann Null, wenn entweder die Regressionskoeffizienten b^m und b^w übereinstimmen (keine Interaktion im Sinne der Regressionsanalyse) oder der Mittelwertsunterschied der unabhängigen Variablen Null ist.

Summa summarum klärt unsere Analyse die folgenden Punkte:

- Die Schätzungen d_1 und g_2 nach dem Duncan-Verfahren erweisen sich als Spezialfall des allgemeinen Modells, wenn die geschlechtsspezifischen Entlohnungsmuster als Regressionsgleichungen dargestellt werden.
- Dadurch erfahren die Größen $d_1 = ID$ und $g_2 = ZD$ eine neue Deutung im Sinne unserer Definitionen als "durch Diskriminierung entgangenes Pro-Kopf-Einkommen".
- Es erfolgt eine klare Bestimmung, welche Diskriminierungsschätzung zu wählen ist. Sinnvoll interpretierbar sind die Größen d_1 und g_2 , während die Komponenten g_1 und d_2 zusätzlich den Interaktionseffekt einschließen und zwei verschiedene Arten von Effekten vermischen.

c) Regressionsanalyse unter Einschluß von qualitativen Variablen

Beim Duncan-Verfahren wird davon ausgegangen, daß alle Variablen quantitativ meßbar sind. Wenn jedoch - wie bei der Regressionsanalyse mit sogenannten Dummy-Variablen - neben quantitativen unabhängigen auch qualitative unabhängige Variablen in der Einkommensfunktion auftreten, dann ist eine Erweiterung des regressionsanalytischen Diskriminierungsmodells erforderlich.

Als Beispiel gehen wir von einer Einkommensfunktion mit der quantitativen Variable "Bildung" und der qualitativen Variable "Leistungsgruppen" aus. Es existieren p Leistungsgruppen, wobei der Index i die Leistungsgruppe bezeichnet. Es läßt sich dann für den Spezialfall der Dummy-Regression nachweisen, daß die folgenden Formeln aus unserem Modell ableitbar sind (zur Ableitung siehe den Anhang):

$$(29) \quad ID = \sum_{i=1}^P f_i^w (\hat{y}_i^{m*} - \bar{y}_i^w) = \sum_{i=1}^P f_i^w \hat{y}_i^{m*} - \bar{y}^w$$

$$(30) \quad ZD = \sum_{i=1}^P f_i^m \hat{y}_i^{w*} - \bar{y}^w$$

Die Formeln (29) und (30) sehen den Ausdrücken (3) und (4) sehr ähnlich. Der Unterschied besteht aber darin, daß in (29) und (30) von "bereinigten" Einkommenschätzungen ausgegangen wird, bei denen der Einfluß der quantitativen unabhängigen Variablen Berücksichtigung findet. Diese bereinigten Schätzungen lauten in unserem Beispiel (x = Bildung):

$$(31) \quad \hat{y}_i^{m*} = b_i^m \bar{x}_i^w + c_i^m$$

\hat{y}_i^{m*} ist das hypothetische Einkommen einer Frau in Leistungskategorie i mit der durchschnittlichen Bildung der Frauen in Kategorie i , wenn das männliche Entlohnungsmuster angewandt wird (wenn also nur die interne Diskriminierung in Kategorie i entfällt): Entsprechend ist \hat{y}_i^{w*} das weibliche Einkommen in Kategorie i , wenn die Bildungs-Zugangschancen der Frauen denen der Männer angepaßt werden. Dieser Wert errechnet sich nach Formel (32):

$$(32) \quad \hat{y}_i^{w*} = b_i^w \bar{x}_i^m + c_i^w$$

Werden a priori für alle Leistungsgruppen die gleichen Regressionskoeffizienten unterstellt, dann entfällt der Index i bei den b -Koeffizienten.

Das Verfahren ist also zweistufig. Zunächst werden die Effekte der quantitativen Einkommensfaktoren innerhalb jeder Leistungskategorie berücksichtigt. Das Resultat sind "bereinigte" Einkommenschätzungen für jede Kategorie. In der zweiten Stufe wird dann ID und ZD wie im einfachen Fall (Formeln (3) und (4)) berechnet. Daß die vorgeschlagene Prozedur nicht nur intuitiv einleuchten mag, sondern sich auch genauer rechtfertigen läßt, wird im Anhang gezeigt.

Noch ein abschließender Hinweis sei gegeben: Das allgemeine Modell (Formeln (12) und (13)) bietet zwar den Vorteil, daß weder spezifische Ein-

kommensfunktionen (z.B. Linearitätsannahme) noch Annahmen über das Skalenniveau erforderlich sind. Wenn jedoch einige oder alle einkommensbestimmenden Variablen mindestens intervallskaliert sind, so empfiehlt sich zumeist die Anwendung der Regressionsanalyse. Der Grund ist vor allem darin zu sehen, daß bei dem allgemeineren tabellenanalytischen Verfahren bei Einbeziehung mehrerer Variablen die Fallzahlen vieler Ausprägungskombinationen gering oder gar Null sein werden, während dies für die Regressionsanalyse (die ja eine Modellstruktur in Form der Einkommensfunktion unterstellt) kein Hindernis darstellt.

Der Chancenungleichheitsindex von Fend und Mitarbeitern

Zur Messung der Zugangschancendiskriminierung wurde ein hauptsächlich im Bildungsbereich angewandter Chancenungleichheitsindex (CUG-Maß) von Fend und Mitarbeitern (Fend u.a. 1976: 214-218) entwickelt. Dieser Index erscheint bei Problemstellungen mit nominaler oder ordinaler "Belohnungsvariable" geeignet zu sein, weniger jedoch, wenn die in Frage stehende abhängige Variable wie bei der Einkommensdiskriminierung quantitativ meßbar ist.

Im Vergleich mit dem ZD-Maß schneidet der CUG-Index schlechter ab, wie aus der folgenden Formel zur Berechnung des CUG-Wertes im Spezialfall der Einkommensdiskriminierung zu ersehen ist.

$$(33) \quad \text{CUG} = 50 \sum |f_i^m - f_i^w|$$

Der Index gibt den Prozentanteil der bestehenden Chancenungleichheit an der maximal möglichen Chancenungleichheit an. Für die Angestelltendaten des Jahres 1980 hat er den Wert 40.

Wie aus Formel (33) hervorgeht, werden jedoch nur die absoluten Differenzen der Zugangschancen für jede Position summiert, wobei die Einkommensunterschiede zwischen den Positionen keine Rolle spielen. Ausdruck (6) dagegen benützt zur Berechnung des ZD-Werts das Einkommen der Frauen als Gewichtungsfaktor⁸.

4. Trends der Diskriminierung weiblicher Angestellter in der Industrie

Wie hoch ist der Anteil der externen und wie hoch der Anteil der internen Diskriminierung an der gesamten Verdienstspanne, wenn als einkommensbestimmender Faktor bei Industrie-Angestellten nur die Leistungsgruppen zugrunde gelegt werden? Weiterhin: Wie gestaltet sich die Entwicklung im Zeitablauf für die vergangene Dekade 1970 bis 1980? Eine Antwort auf diese Fragen kann auf der Grundlage der Tabellen 1 und 2 gegeben werden. Tabelle 1 informiert über die durchschnittlichen Bruttoverdienste nach Leistungsgruppe, Geschlecht und Jahreszahl.

Tabelle 1: Bruttoverdienste und Verdienstspanne bei Industrie-Angestellten nach Leistungsgruppen und Geschlecht

Jahr	Bruttomonatsverdienste \bar{y}_i^w der weibl. Angestellten (DM)					abs.Differenz zu den Ver- diensten ($\bar{y}_i^m - \bar{y}_i^w$) der männl. Angestellten (DM)				
	Leistungsgruppe					Leistungsgruppe				
	II	III	IV	V	alle \bar{y}^w	II	III	IV	V	alle $(\bar{y}^m - \bar{y}^w)$
1970	1581	1202	894	714	1009	428	321	276	246	603
1975	2606	1961	1499	1194	1732	534	472	395	370	901
1976	2793	2099	1604	1286	1859	561	496	415	362	959
1977	2981	2255	1726	1393	2004	604	519	436	384	1017
1978	3144	2384	1829	1484	2127	638	547	453	414	1057
1979	3322	2515	1933	1576	2256	700	589	485	439	1143
1980	3552	2688	2067	1695	2416	764	628	514	457	1232

Quelle: Amtliche Verdienststatistik des Statistischen Bundesamtes Wiesbaden, Fachserie 16, Reihe 2.2 (siehe auch Anmerkung 2)

Weiterhin benötigen wir noch Angaben zu den geschlechtsspezifischen Positionsverteilungen, die aus Tabelle 2 hervorgehen:

Tabelle 2: Verteilung auf die Leistungsgruppen nach Geschlecht

Jahr	Männer (f_i^m)					Frauen (f_i^w)				
	Leistungsgruppe					Leistungsgruppe				
	II	III	IV	V	alle	II	III	IV	V	alle
1970	32	51	16	1	100	5	34	50	11	100
1975	38	50	11	1	100	6	40	47	7	100
1976	39	50	10	1	100	7	40	46	7	100
1977	39	50	10	1	100	7	41	46	6	100
1978	39	50	10	1	100	7	42	45	6	100
1979	40	49	10	1	100	7	42	45	6	100
1980	41	49	9	1	100	8	42	44	6	100

Quelle: siehe Tabelle 1

Gemäß Tabelle 1 läßt sich für alle Positionen interne Diskriminierung nachweisen, da alle Differenzen positiv sind. Der absolute Einkommensabstand hat sich von 1970 bis 1980 mehr als verdoppelt, wobei allerdings der Preisanstieg noch zu berücksichtigen ist. Immerhin nehmen die Verdienstspannen absolut auch dann zu, wenn die Preise des Jahres 1980 zugrunde gelegt werden. Der prozentuale Mehrverdienst der Männer geht dagegen leicht zurück.

Betrachten wir nun die Verteilung auf die Leistungsgruppen anhand Tabelle 2. Eine erste Einschätzung läßt erkennen:

- Auch 1980 ist die Chance der Männer, die oberste Position (Leistungsgruppe II) einzunehmen, fünfmal so groß wie die Chance der Frauen. Genau spiegelbildlich verhält es sich bei der untersten Leistungsgruppe

pe. Hier ist die Chance der Frauen sechsmal so groß wie die der Männer.

- Zwischen 1970 und 1980 erfolgte eine Qualifikationsverschiebung nach oben für Männer und Frauen. In den Kategorien II und III sind 1980 11% mehr Frauen und 7% mehr Männer als 1970 beschäftigt.
- Im einzelnen betrachtet, holen die Frauen allerdings vorwiegend in Kategorie III auf, während sich in der obersten Leistungsgruppe II der Chancenabstand zwischen Männern und Frauen noch vergrößert.
- Die Chancenabstände kompensieren sich weitgehend. In Kategorie II werden sie ungünstiger, in Kategorie III günstiger, in Kategorie IV bleiben sie ziemlich gleich und in Kategorie V wiederum verringern sie sich. Wegen dieser Kompensationseffekte zeigt sich insgesamt nur eine leichte Tendenz zur Verminderung der Chancendifferenzen, wenn das CUG-Maß herangezogen wird (Tabelle 3).

Tabelle 3: CUG-Maß für Zugangschancen zu Leistungsgruppen nach Geschlecht

Jahr	CUG
1970	44
1975	42
1976	42
1977	41
1978	40
1979	40
1980	40

Wie im vorhergehenden Abschnitt erläutert, berücksichtigt das CUG-Maß jedoch nicht die Einkommensabstände zwischen den Leistungsgruppen. Es kann daher nur als grober Indikator der Zugangschancendiskriminierung angesehen werden.

Genaue Auskünfte über die quantitative Höhe des Einkommensverlusts der Frauen aufgrund ungünstigerer Zugangschancen (ZD) und interner Diskriminierung gibt (ID) Tabelle 4.

Tabelle 4: Einkommensverlust durch externe (ZD) und interne (ID) Diskriminierung in DM

Jahr	Verdienstspanne ($\bar{y}^m - \bar{y}^w$)	Absolute Werte in DM			Prozentuale Anteile		
		ZD	ID	IE	ZD%	ID%	IE%
1970	603	256	296	51	42,4	49,1	8,5
1975	901	419	432	50	46,5	48,0	5,5
1976	959	449	454	56	46,8	47,3	5,8
1977	1017	466	479	72	45,8	47,1	7,1
1978	1057	482	503	72	45,6	47,6	6,8
1979	1143	517	541	85	45,2	47,3	7,4
1980	1232	552	578	102	44,8	47,0	8,3

Existierte keine Zugangschancendiskriminierung, dann hätten Frauen 1970 im Durchschnitt 256 DM mehr erhalten, und im Jahre 1980 bezifferte sich der Mehrgewinn auf 552 DM. Läge dagegen keine interne Diskriminierung vor, dann hätten Frauen 1970 DM 296 mehr in der Lohntüte und 1980 DM 578. Wären beide Diskriminierungsarten hypothetisch beseitigt, dann ist noch der "Interaktionseffekt" hinzuzuzählen, d.h. das durchschnittliche Frauen-Einkommen entspräche dann dem durchschnittlichen Männer-Einkommen.

Dramatische Veränderungen sowohl in Hinblick auf die Höhe als auch die Zusammensetzung der Diskriminierung zeigen sich in der betrachteten Dekade nicht. Vielmehr sind die Verhältnisse relativ stabil mit einer leichten Tendenz zum Abbau der prozentualen Diskriminierung: Der aufgrund ungünstigerer Zugangschancen entgangene Verdienst beträgt 1970 25% des Frauen-Einkommens und 1980 23%. Bei der internen Diskriminierung sind es 1970 29% und 1980 24%.

Auch die Anteile von ZD und ID an der Verdienstspanne sind in dem betrachteten Zeitraum ziemlich konstant geblieben. Hier läßt sich allenfalls

ein leichter Trend zur Verminderung interner Diskriminierung ausmachen. Es sei abschließend nochmals darauf hingewiesen, daß wir bei dieser Analyse den Begriff "Diskriminierung" verwendet haben, obwohl nur die unabhängige Variable "Leistungsgruppen" kontrolliert wurde. Die erwerbstätigen Frauen und Männer unterscheiden sich - wie gesagt - auch durch eine Reihe anderer einkommensbestimmender Merkmale. So mögen die Zugangschancen der Frauen z.B. deshalb geringer sein, weil Ausbildungsunterschiede, Berufserfahrungsdefizite usw. vorliegen. Dennoch erscheint es berechtigt, von Diskriminierung zu sprechen, wenn diese als Ausdruck der Kumulation von Benachteiligungen auf früheren Stufen des Lebenszyklus (neben einer "reinen" Geschlechtsdiskriminierung) aufgefaßt wird (vgl. auch Duncan 1968).

Auch wenn gemäß dieser Überlegungen ZD und ID z.T. auf kumulierte frühere Benachteiligungen zurückführbar sind, so wäre es doch von großem Interesse, die Stärke der einzelnen Komponenten von ZD und ID zu identifizieren. Es fragt sich also, welcher Anteil der Verdienstspanne durch Ausbildungs Nachteile, welcher durch Berufserfahrungs Nachteile usw. zu erklären ist. Zur Beantwortung der erwähnten Fragen sind jedoch Individualdaten zur Einkommenssituation und zu den einkommensbestimmenden Variablen erforderlich. Die groben Aggregate der amtlichen Statistik gestatten leider keine derart verfeinerten Analysen. Immerhin zeigen erste Untersuchungsergebnisse auf der Basis von Individualdaten, daß auch bei der Berücksichtigung von Bildung, Ausbildung und Berufserfahrung die Einkommensabstände zwischen Männern und Frauen keineswegs in hohem Grade reduziert werden. Offensichtlich existiert auch bei Kontrolle wichtiger Einkommensfaktoren ein hohes Ausmaß an interner Diskriminierung, wie sich im folgenden Abschnitt zeigen wird.

5. Eine multivariate Analyse der Einkommensunterschiede

Auf der Basis von Umfragedaten des ALLBUS 1980 wollen wir die Einkommensunterschiede bei der Gruppe von Angestellten und Beamten, aufgliedert nach vier "Leistungskategorien", etwas genauer betrachten⁹.

In Tabelle 5 sind zunächst die Einkommensmittelwerte und die relativen Häufigkeiten der Kategorienbesetzung für Männer und Frauen aufgeführt¹⁰.

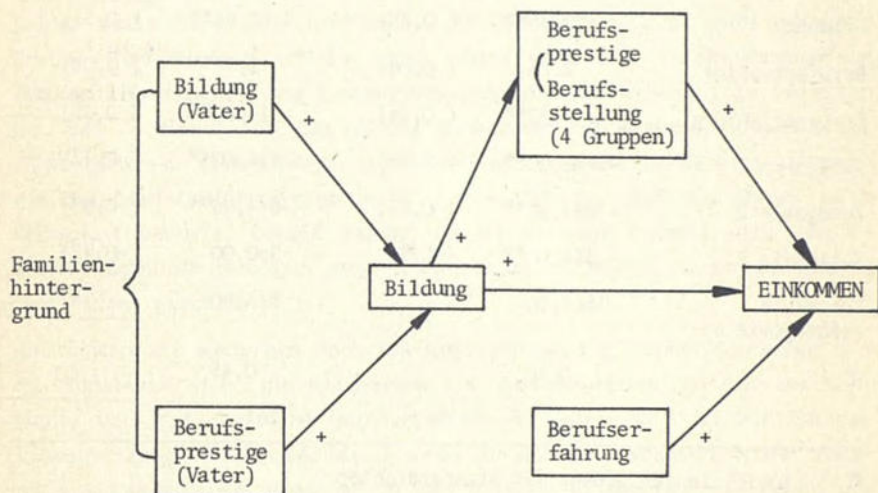
Tabelle 5: Netto - Einkommen (Mittelwerte) und Häufigkeitsverteilung nach Leistungskategorien und Geschlecht für Angestellte und Beamte in DM

Kategorien	Einkommen			Relative Häufigkeit		
	Männer	Frauen	Differenz	Männer	Frauen	Differenz
1	1588	1183	405	0,12	0,22	-0,10
2	2060	1521	539	0,43	0,57	-0,13
3	2676	1770	906	0,37	0,19	0,18
4	3529	2575	954	0,08	0,02	0,06
insgesamt	2355	1518	837	1,00	1,00	1,00
ID = 588	ID in % = 70		$N^m = 284$			
ZD = 145	ZD in % = 17		$N^w = 161$			
IE = 104	IE in % = 12					

Bei einem qualitativen Vergleich zeigen sich die gleichen Tendenzen wie bei den Angestellten in der Industrie. Es liegt sowohl Zugangschancendiskriminierung als auch interne Diskriminierung vor¹¹. Die Existenz von ZD ist auf einen Blick erkennbar: In Spalte 7 der Tabelle sind die Differenzen negativ für die unteren und positiv für die oberen Kategorien. Ferner ist ID gegeben, da in jeder Kategorie Männer mehr verdienen als Frauen. Quantitativ gesehen entgehen den Frauen aufgrund interner Diskriminierung durchschnittlich 588 DM und aufgrund von Zugangschancendiskriminierung DM 145. Es fragt sich nun, ob bei Berücksichtigung weiterer Variablen wie Bildung und Dauer der Erwerbstätigkeit das Ausmaß der internen Diskriminierung reduziert wird. Ein geeignetes Verfahren zur

Kontrolle der sozioökonomischen Hintergrundmerkmale ist die Regressionsanalyse. Zur Erklärung des Nettoeinkommens spezifizieren wir das in Abbildung 1 gezeigte Pfadmodell¹².

Abbildung 1: Regressionsmodell zur Erklärung des Netto - Einkommens durch sozioökonomische Variablen des Lebenszyklus



+: positive Beziehung erwartet

Da an dieser Stelle die Erklärung des Netto-Einkommens von Interesse ist, konzentrieren wir uns im folgenden nur auf die Regressionsgleichung für das Einkommen¹³.

Aus Tabelle 6 lassen sich die Ergebnisse der Regressionsanalyse entnehmen.

Tabelle 6: Regressionskoeffizienten für abhängige Variable Netto - Einkommen in DM

	Männer		Frauen	
Bildung Vater	- 7,85	(-0,02)***	38,43*	(0,13)
Berufsprestige Vater	3,85	(0,05)	- 2,10	(-0,05)
Bildung	73,70**	(0,25)	54,64**	(0,31)
Berufsprestige	2,73	(0,04)	3,97*	(0,08)
Berufserfahrung	19,32**	(0,26)	15,00**	(0,35)
Kategorie 1	-1430,00**	(-0,55)	-894,00**	(-0,77)
Kategorie 2	-1095,00**	(-0,65)	-614,00**	(-0,63)
Kategorie 3	- 654,00**	(-0,38)	-560,00	(-0,45)
Konstante (=Kategorie 4)	1663,00		823,00	
R ²	0,46		0,45	
N	284		161	

* Koeffizienten größer als Standardfehler

** Koeffizienten größer als zweifacher Standardfehler

*** standardisierte Koeffizienten in Klammern

Der unstandardisierte Regressionskoeffizient gibt darüber Auskunft, in welchem Umfang das Einkommen steigt oder sinkt, wenn die jeweilige unabhängige Variable um eine Einheit zunimmt. Betrachten wir, um dies anschaulicher zu erläutern, das Beispiel der unabhängigen Variable Berufsprestige bei der Frauen-Stichprobe. Unterscheiden sich zwei Gruppen von Frauen in bezug auf die Ausprägungen der unabhängigen Variablen nur bei dem Merkmal Berufsprestige derart, daß die eine Gruppe von Befragten ein um einen Skalenpunkt höheres Berufsprestige aufweist, so liegt für diese Gruppe das Einkommen im Mittel um 3,97 DM höher. Bei einem Unterschied von einem Jahr Schulbildung (und ansonsten gleichen Merkmalsausprägungen) beträgt der erwartete Einkommensunterschied 54,64 DM.

Häufig werden die Koeffizienten auch in folgender Weise interpretiert: Wenn eine Frau in ein zusätzliches Bildungsjahr investiert, dann erhöht sich ihr erwartetes Einkommen um 54,64 DM. Ist diese Interpretation nicht nur hypothetisch, sondern als Wirkungsprognose zu verstehen, dann sind allerdings zwei Punkte zu berücksichtigen. Erstens muß es sich um einen kausalen Zusammenhang handeln, was bei Modellen des in Abbildung 1 dargestellten Typs meist angenommen wird. Zweitens ist zu berücksichtigen, daß die "Politikempfehlung" Erhöhung von Bildungsinvestitionen eine Dynamik auslösen kann, deren Resultat mit den Modellerwartungen in keiner Weise mehr übereinstimmen muß. Verlängern z.B. viele oder alle Frauen ihre Bildung um ein Jahr, dann wird sich ihr Einkommen bei starrer Arbeitsplatz- und Einkommensstruktur kaum erhöhen. Zu vermuten ist, daß in diesem Fall das gleiche Einkommen bei allgemein angehobenem Bildungsniveau erzielt wird. Die erwähnte Dynamik könnte also - ähnlich wie bei "self-fulfilling-prophecies" - dazu führen, daß das Modell seine Gültigkeit verliert. Dieser Aspekt sei als Warnung betont, allzu weitreichende Schlußfolgerungen aus einem statischen Modell zu ziehen, wie es des öfteren geschieht.

Betrachten wir weiterhin noch die Koeffizienten der Dummy-Variablen und die Konstante, d.h. die Kategorien 1-4. Ausgangspunkt ist hier die Konstante von z.B. 1663 DM bei den Männern. Dieser Wert ist zur Einkommenschätzung hinzuzuaddieren, wenn es sich um Personen in der Kategorie 4 handelt. Bei Personen in den übrigen Kategorien muß davon der jeweilige Dummy-Koeffizient abgezogen werden, um den fixen "Sockelbetrag" zu errechnen. Für Kategorie 1 erhalten wir somit einen Sockelbetrag von $1663 - 1430 = 233$ DM. In Kategorie 2 sind es 568 DM, in Kategorie 3 1009 DM und in Kategorie 4 wie gesagt 1663 DM. Der Zusammenhang mit dem Einkommen ist entsprechend den Erwartungen positiv. Die jeweiligen Sockelbeträge sind wie Konstanten zur Regressionsschätzung zu addieren.

Nach diesen Erläuterungen seien die wichtigsten Ergebnisse in Stichworten präsentiert:

- Insbesondere bei Männern üben die familiären Hintergrundmerkmale so gut wie keinen direkten Effekt auf die Einkommenshöhe aus. Bei Frauen zeigt sich für die Bildung des Vaters ein schwacher Einfluß. Der familiäre Hintergrund verschafft sich hauptsächlich indirekt (z.B. über die Bildung) Geltung.

- Bei Berücksichtigung der Berufsstellung stellt das Berufsprestige keinen erklärungskräftigen Einkommensprädiktor dar.
- Von der Berufsstellung abgesehen, spielen die Variablen Bildung (einschließlich Berufsausbildung) und Berufserfahrung die Hauptrolle bei der Erklärung der Einkommenshöhe.
- Auffallende Geschlechtsunterschiede sind bei den Koeffizienten der Ausbildung erkennbar. Grob gesprochen verhilft ein Jahr zusätzliche Bildung einem Mann zu Mehreinkünften von ca. 74 DM, während es bei Frauen etwa 55 DM sind. Ein Jahr zusätzliche Berufserfahrung schlägt bei Männern mit 19 DM und bei Frauen mit 15 DM zu Buche.

Zur Beantwortung der eingangs gestellten Frage nach der Höhe der internen Diskriminierung bei Kontrolle der sozioökonomischen Hintergrundmerkmale benötigen wir noch die Mittelwerte der unabhängigen Variablen für die vier Kategorien getrennt nach Geschlecht. Tabelle 7 informiert über die Mittelwerte.

Tabelle 7: Mittelwerte der unabhängigen Variablen nach Leistungskategorien und Geschlecht

M ä n n e r						
Kate- gorien	Bildung Vater	Berufs- prestige Vater	Bildung	Berufs- prestige	Berufs- erfah- rung	N
1	8,33	37,91	9,98	38,82	22,45	33
2	8,43	39,50	10,94	44,79	24,75	123
3	9,34	42,88	13,38	50,46	23,34	104
4	9,96	50,71	16,23	57,63	20,50	24
F r a u e n						
1	8,56	39,72	10,36	38,75	19,22	36
2	8,79	41,13	11,66	47,80	15,55	91
3	9,60	45,17	13,35	49,37	20,50	30
4	10,75	47,75	18,00	60,00	14,50	4

Mit den Angaben in Tabelle 6 und Tabelle 7 kann nun das Ausmaß interner Diskriminierung geschätzt werden. Dazu stellen wir die hypothetische Frage: Welches Einkommen erhält eine Frau im Mittel in der Leistungskategorie i , wenn sie die durchschnittlichen Voraussetzungen bezüglich Bildung, Berufserfahrung etc. bezogen auf die Frauen in der jeweiligen Kategorie mitbrächte, aber ihre Einkünfte nach dem Entlohnungsmuster der Männer erhielte? Dieser Schätzwert (\hat{y}_i^{m*}) errechnet sich, indem die Mittelwerte der Frauen in Kategorie i in die Regressionsgleichung für Männer eingesetzt werden. Das durchschnittliche Ausmaß der internen Diskriminierung bei Berücksichtigung der genannten Einkommensfaktoren (ID) kann dann mit Formel (29) bestimmt werden. Die auf diese Weise gewonnenen Ergebnisse zeigt Tabelle 8.

Tabelle 8: Schätzung der internen Diskriminierung bei Kontrolle der sozio-ökonomischen Variablen

	Kategorie			
	1	2	3	4
Einkommen Männer \bar{y}_i^m	1588	2060	2676	3529
Hypothetisches Einkommen der Frauen \hat{y}_i^{m*}	1559	1947	2623	3533
Tatsächliches Einkommen der Frauen \bar{y}_i^w	1183	1521	1770	2575
Diskriminierungslücke $(\hat{y}_i^{m*} - \bar{y}_i^w)$	376	426	843	958
Relative Häufigkeit f_i^w	0,22	0,57	0,19	0,02
$f_i^w (\hat{y}_i^{m*} - \bar{y}_i^w)$	83	243	160	19

Demnach beträgt der Einkommensverlust aufgrund interner Diskriminierung pro Kopf der Frauen 505 DM (Summe der letzten Zeile in Tabelle 8). Der Anteil der internen Diskriminierung an der globalen Einkommensdifferenz von 837 DM (Tabelle 5) beträgt somit 60%. Die Verminderung von ID gegenüber der Analyse von Tabelle 5 um 10% ist in erster Linie durch die Berücksichtigung der geringeren Berufserfahrung der Frauen zu erklären.

Betrachten wir abschließend noch die Zugangschancendiskriminierung anhand von Tabelle 9.

Tabelle 9: Schätzung der Zugangschancendiskriminierung bei Kontrolle der sozioökonomischen Variablen

	Kategorie			
	1	2	3	4
Hypothetisches Einkommen der Frauen \hat{y}_i^{w*}	1206	1597	1813	2522
Zugangschancen der Männer f_i^m	0,12	0,43	0,37	0,08
Produkt $f_i^m \cdot \hat{y}_i^{w*}$	145	687	671	202

Gemäß Formel (30) errechnet sich für ZD ein Wert von:

$$ZD = \sum_{i=1}^p f_i^m \hat{y}_i^{w*} - \bar{y}^w = 1704 - 1518 = 186$$

Nach dieser Rechnung entgehen den Frauen im Durchschnitt 186 DM aufgrund der ungünstigeren Verteilung bei den einkommensbestimmenden

Faktoren. Der Hauptanteil davon ist auf die Leistungsgruppen-Verteilung und ein kleiner Anteil auf die geringere Berufserfahrung der Frauen zurückzuführen (vgl. Tabelle 7). Die quantitative Zerlegung des Einkommensabstands zwischen Männern und Frauen in die zwei Komponenten ID und ZD sowie deren Wechselwirkung ist Tabelle 10 zu entnehmen.

Tabelle 10: Zerlegung des Einkommensabstandes bei Kontrolle der sozio-ökonomischen Variablen in DM

	ID	ZD	IE	$(\bar{Y}^m - \bar{Y}^w)$
absolut	505	186	146	837
prozentual	60,3	22,2	17,4	100

Wenn also weibliche Angestellte und Beamte bei gleicher Ausbildung, gleicher Berufserfahrung, gleichem Berufsprestige, gleichem familiären Hintergrund, bei Ganztagsarbeit und gleicher Leistungskategorie ein genauso hohes Einkommen erhielten wie die Männer mit genau diesen Merkmalen, dann würden sie - bezogen auf das Jahr 1980 - durchschnittlich 505 DM höhere monatliche Nettoeinkünfte erhalten. Dies sind 33% mehr als das tatsächliche Einkommen von 1518 DM. Wenn ferner hinsichtlich der sozio-ökonomischen Variablen und der Leistungsgruppen die Verteilung der Frauen der männlichen Verteilung entspräche, wenn also keine Zugangschancendiskriminierung zwischen den beiden Geschlechtern existierte, dann wäre das hypothetische Pro-Kopf-Einkommen der Frauen 186 DM oder 12% höher als das tatsächliche Einkommen. Der größte Anteil an der globalen Einkommenslücke geht also auf das Konto interner Einkommensdiskriminierung. Insgesamt zeigt sich - im Einklang mit internationalen Untersuchungen - daß auch bei der Berücksichtigung zentraler Faktoren der Einkommensbildung der Anteil der von uns als interne Diskriminierung bezeichneten Ungleichheitskomponente nur wenig vermindert wird¹⁴.

Zusammenfassung

Die Einkommensunterschiede zwischen Männern und Frauen lassen sich auf zwei Mechanismen der Diskriminierung zurückführen. Zum einen ist die Verteilung der Frauen auf die einkommensbestimmenden Positionen ungünstiger als die der Männer und zum zweiten erzielen Frauen auch dann, wenn sie die gleichen Voraussetzungen mitbringen wie Männer, im Mittel geringere Einkünfte.

In dem vorliegenden Aufsatz wurde zunächst ein Verfahren entwickelt, das die Berechnung der quantitativen Höhe der Diskriminierungskomponenten gestattet. Ein besonderer Abschnitt war dem Zusammenhang zwischen dem vorgeschlagenen Modell und dem Diskriminierungsmodell von Duncan gewidmet. Es zeigte sich, daß das Duncan-Modell unter bestimmten Bedingungen als Spezialfall des allgemeinen Diskriminierungsmodells gedeutet werden kann. Erweiterungen des Modells bezogen sich ferner auf Regressionsschätzungen unter Einschluß qualitativer Variablen.

Anhand amtlicher statistischer Daten aus der Bundesrepublik wurden Diskriminierungsschätzungen für weibliche Angestellte in der Industrie für den Zeitraum 1970 bis 1980 präsentiert. Die amtlichen Daten erlauben allerdings nur eine grobe Analyse, da zentrale Hintergrundmerkmale wie Bildung, familiäre Ressourcen usf. unkontrolliert bleiben. Um diesem Mangel abzuhelpen, wurde ein multivariates Pfadmodell spezifiziert, dessen Koeffizienten an Umfragedaten des ALLBUS 1980 geschätzt wurden. Diskriminierungsberechnungen ergaben, daß bei Angestellten und Beamten das Ausmaß der "direkten" Einkommensdiskriminierung auch dann beträchtlich ist, wenn den sozioökonomischen Hintergrundmerkmalen Rechnung getragen wird.

Anhang

ID und ZD für den Fall der Dummy-Regression

Wir wollen zeigen, daß die Gleichungen (29) und (30) aus dem allgemeinen Modell ableitbar sind, wenn in (12) und (13) anstelle der Einkommens-Mittelwerte die Regressionsschätzungen eingesetzt werden. j ist der Index

für die Ausprägung der quantitativen Variable, i bezeichnet die Kategorie der qualitativen Variable. Summiert wird stets von $j=1$ bis n (Zahl der Ausprägungen) und $i=1$ bis p (Zahl der Kategorien). f_{ij}^w bzw. f_{ij}^m sind die relativen Häufigkeiten der zweidimensionalen weiblichen bzw. männlichen Verteilung. \hat{y}_{ij}^m bzw. \hat{y}_{ij}^w sind die Regressionsschätzungen für jede Ausprägungskombination der unabhängigen Variablen. Dann gilt gemäß (12):

$$ID = \sum_i \sum_j f_{ij}^w \hat{y}_{ij}^m - \bar{Y}^w$$

Bei einer quantitativen unabhängigen Variablen x lautet die Schätzgleichung:

$$\hat{y}_{ij}^m = b_i^m x_{ij} + c_i^m \quad \text{für } i=1, \dots, p$$

Eingesetzt in obige Gleichung erhält man:

$$ID = \sum_i b_i^m \left[\sum_j (f_{ij}^w x_{ij}) + c_i^m \sum_j f_{ij}^w \right] - \bar{Y}^w$$

Nun gelten folgende Beziehungen: $\sum_j (f_{ij}^w / f_i^w) x_{ij} = \bar{X}_i^w$ und $\sum_j (f_{ij}^w / f_i^w) = 1$ für $i=1, \dots, p$. Daher kann die ID-Gleichung weiter vereinfacht werden:

$$ID = \sum_i f_i^w (b_i^m \bar{X}_i^w + c_i^m) - \bar{Y}^w$$

Schreiben wir für den Ausdruck in der runden Klammer \hat{y}_i^{m*} , dann ergibt sich Gleichung (29).

Für ZD gehen wir in analoger Weise vor. Gemäß (13) gilt:

$$ZD = \sum_i \sum_j f_{ij}^m \hat{y}_{ij}^w - \bar{Y}^w$$

$$\text{mit } \hat{y}_{ij}^w = b_i^w x_{ij} + c_i^w$$

Aus den beiden Gleichungen folgt:

$$ZD = \sum_i \left[b_i^w \sum_j (f_{ij}^m x_{ij}) + c_i^w \sum_j f_{ij}^m \right] - \bar{Y}^w$$

Unter Berücksichtigung von $\sum_j (f_{ij}^m / f_i^m) x_{ij} = \bar{X}_i^m$ und $\sum_j (f_{ij}^m / f_i^m) = 1$ folgt:

$$ZD = \sum_i f_i^m (b_i^w \bar{X}_i^m + c_i^w) - \bar{Y}^w$$

Der Ausdruck in der runden Klammer ist der Regressionsschätzwert \hat{y}_i^{w*} , so daß sich die ZD-Gleichung zu Gleichung (30) vereinfacht. Alle Überlegungen gelten auch entsprechend für die multivariate Erweiterung auf mehrere quantitative Einkommensdeterminanten.

Anmerkungen

- 1 Für kritische Kommentare zu dieser Arbeit bedanke ich mich herzlichst bei Joseph Christl, Walter Krämer, Karl Ulrich Mayer, Peter Schmidt und Gilg Seeber.
- 2 Für die Zusendung der vom Wirtschafts- und Sozialwissenschaftlichen Institut des DGB (WSI) aufbereiteten Daten bedanke ich mich herzlichst bei Frau Brigitte Stahn-Willig.
Die genaue Definition der von der amtlichen Statistik verwendeten Leistungsgruppen ist der Einleitung der halbjährlich erscheinenden Hefte der Fachserie 16, Reihe 2.2 des Statistischen Bundesamts Wiesbaden zu entnehmen. Grob gesprochen handelt es sich bei Leistungsgruppe II um kaufmännische und technische Angestellte mit besonderen Erfahrungen in verantwortlicher Tätigkeit mit eingeschränkter Dispositionsbefugnis, die Angestellte anderer Tätigkeitsgruppen unterweisen. Leistungsgruppe III umfaßt Angestellte mit mehrjähriger Berufserfahrung oder besonderen Fachkenntnissen, die jedoch in der Regel keine Verantwortung für die Tätigkeit anderer tragen. In Gruppe IV befinden sich Angestellte ohne eigene Entscheidungsbefugnis mit einfacher Tätigkeit, aber abgeschlossener Berufsausbildung oder mehrjähriger Berufserfahrung. Angestellte in Leistungsgruppe V üben einfache schematische Tätigkeiten aus, die keine Berufsausbildung erfordern.
- 3 Dies ist dann der Fall, wenn die Einkommens- und Häufigkeitsdifferenzen $(y_i^m - y_i^w)$ bzw. $(f_i^m - f_i^w)$ positiv korreliert sind. Den Nachweis dieses Satzes verdanke ich Walter Krämer.
- 4 Sind einige erklärende Variablen auf Intervallskalenniveau meßbar, andere dagegen nur auf qualitativem Niveau, so kann das statistische Verfahren der Kovarianzanalyse oder Dummy-Regression benützt werden (dazu weiter unten).

- 5 Zur Darstellung und Kritik des Verfahrens siehe auch Diekmann (1981: 64-76).
- 6 Wenn strukturelle Privilegien und Privilegien bezüglich der unabhängigen Variablen existieren, gilt:

$$b^m > b^w \text{ und } \bar{X}^m > \bar{X}^w.$$

In diesem Fall läßt sich nachweisen, daß d_2 größer als d_1 ist. Wenn $b^m = b^w$, ist auch $d_2 = d_1$.

- 7 Nach Drucklegung dieser Arbeit hat mich Robert M. Hauser, Wisconsin, dankenswerterweise auf einen Artikel von Winsborough und Dickinson (1971) aufmerksam gemacht. Die Autoren argumentieren für die gleiche Zerlegung der Mittelwertdifferenz wie in dem vorliegenden Aufsatz. Gleichung 4 in Winsborough und Dickinson entspricht (27), Gleichung 5 ist mit (26) identisch und der Interaktionseffekt (Gleichung (6)) stimmt exakt mit der Formel für IE in der Form von S.330 oben überein. Winsborough und Dickinson wenden ihr Modell auf die Analyse der Einkommensunterschiede zwischen Weißen und Farbigen an.
- 8 Diese Kritik trifft auch auf Anwendungen des CUG-Maßes im Bildungsbereich zu, wenn die "ChancenvARIABLE" z.B. die Zahl der besuchten Schuljahre ist.
- 9 Der ALLBUS 1980 ist eine haushaltsrepräsentative Umfrage bei 2955 Haushalten in der Bundesrepublik und West-Berlin, die vom Zentrum für Umfragen, Methoden und Analysen in Mannheim (ZUMA) mit Unterstützung der DFG organisiert wurde. Die Daten sind allgemein zugänglich und sehr gut dokumentiert, so daß u.a. der Vorteil erhöhter intersubjektiver Kontrollierbarkeit von Ergebnissen gegeben ist. Bei dem dreistufigen Stichprobenverfahren wurden zunächst (in zwei Stufen) die Haushalte zufällig ausgewählt. Die befragte Person eines jeden gewählten Haushalts wurde sodann nach einem vom Interviewer angewandten Zufallszahlenschlüssel bestimmt. Erhoben wurden sowohl Merkmale des Befragten als auch des Haushalts. Die im folgenden berichteten Ergebnisse beziehen sich nur auf Merkmale des Befragten.
- 10 Bei der Analyse wurden zunächst diejenigen wenigen Fälle ausgesondert, bei denen die Befragten nach Einschätzung des Interviewers unzuverlässige Antworten gegeben haben. (V343 = 3. Für Leser, die an Sekundäranalysen interessiert sind, werden in Klammern die entsprechenden Auswahlkategorien angegeben.) Weiterhin betrachten wir nur ganztags Beschäftigte (V136 = 1). Bei der Umfrage wurde die Berufsstellung der Beamten und diejenige der Angestellten in vier Kategorien aufgegliedert. Wir haben jeweils Beamte und Angestellte der korrespondierenden Leistungskategorien zusammengefaßt, da bei getrennter Analyse von Angestellten und Beamten die Fallzahl zu gering wäre. Die Klassifizierung der Angestellten ist nicht mit der Klassifizierung nach der amtlichen Statistik vergleichbar, die in Teil 4 zugrunde gelegt wurde. Die Leistungskategorien sind wie folgt definiert: Kategorie 1 = Angestellte mit einfacher Tätigkeit (z.B. Verkäufer, Kontorist) und Beamte im einfachen Dienst (V137 = 51 und 40). Kategorie 2 = Angestellte, die schwierigere Aufgaben nach allgemeiner Anweisung selbständig erledigen (z.B. Sachbearbeiter, Buchhalter) und Beamte im mittleren

Dienst (V137 = 52 und 41). Kategorie 3 = Angestellte, die selbständige Leistungen in verantwortungsvoller Tätigkeit erbringen oder begrenzte Verantwortung für die Tätigkeit anderer tragen (z.B. wiss. Mitarbeiter, Prokurist) sowie Beamte im gehobenen Dienst (V137 = 53 und 42). Kategorie 4 = Angestellte mit umfassenden Führungsaufgaben und Entscheidungsbefugnissen (Direktor, Geschäftsführer und Beamte im höheren Dienst (V137 = 54 und 43).

- 11 Die Ergebnisse sind allerdings der quantitativen Höhe nach nicht mit den Befunden bei den Industrie-Angestellten vergleichbar, da es sich im Unterschied zur vorhergehenden Analyse erstens um Angestellte und Beamte und zweitens um Netto-Einkommen handelt.
- 12 Dabei haben die Modelle der Forschungstradition von Blau und Duncan Pate gestanden (siehe z.B. Duncan 1968). Im Unterschied zu Duncan berücksichtigen wir noch die Berufserfahrung und die Berufsstellung, d.h. die vier erwähnten Leistungskategorien, als Dummy-Variablen. Die Variablen wurden in folgender Weise operationalisiert: Einkommen = Netto-Monatseinkommen des Befragten in DM; Bildung des Vaters = Allgemeinbildung in Schuljahren; Bildung des Befragten = Allgemeinbildung und verschiedene Arten der Berufsbildung in Jahren gemäß einem Vorschlag von Helberger (1980); Berufsprestige des Vaters und Berufsprestige des Befragten = Treiman-Skala (siehe Treiman 1979); Berufserfahrung = Alter abzüglich Allgemeinbildung (einschließlich Studium) in Jahren abzüglich sechs Jahre (siehe dazu auch Featherman und Hauser 1976: 465). Diese Operationalisierung der Berufserfahrung ist zweifellos eine sehr grobe Näherung, zumal gerade bei Frauen häufiger auftretende Berufsunterbrechungen unberücksichtigt bleiben.
- 13 Die standardisierten Koeffizienten für die Bildungsgleichung lauten bei Männern: Bildung Vater/Bildung Befragter 0,35, Berufsprestige Vater/Bildung Befragter 0,10. Das R^2 beträgt 0,17. Bei Frauen ergeben sich die Werte: Bildung Vater/Bildung Befragte 0,31, Berufsprestige Vater/Bildung Befragte 0,22, $R^2 = 0,23$.
- 14 Siehe dazu z.B. die Studien von Featherman und Hauser (1976) für die USA und auch die Untersuchung von Swafford (1978) für die Sowjetunion.
Weitere Analysen der Einkommensungleichheiten für Österreich und die BRD mit den Daten des ALLBUS 1980 und des ALLBUS 1982 - sind Gegenstand eines laufenden Forschungsprojekts, das der Verfasser am Institut für Höhere Studien in Wien durchführt.

Literatur

- Christl, Joseph und Michael Wagner, 1982: Die Stellung der Frau in der österreichischen Lohnpyramide. Wirtschaft und Gesellschaft 1/1982: 79-97.
- Diekmann, Andreas, 1981: Sozialindikatoren der Ungleichheit, Chancenungleichheit und Diskriminierung. Anwendung einfacher mathematischer Hilfsmittel in der Ungleichheitsforschung. Forschungsbericht Nr.166 des Instituts für Höhere Studien. Wien: Fachverlag für Wirtschaft und Technik.

- Duncan, Otis D., 1968: Inheritance of Poverty or Inheritance of Race? S.85-110 in: Patrick Moynihan (Hrsg.), On Understanding Poverty. New York: Basic Books.
- Featherman, David L. und Robert M. Hauser, 1976: Sex and Socioeconomic Achievement. *American Sociological Review*, 41:462-463.
- Fend, Helmut, Wolfgang Knörzer, Willibald Nogl, Werner Specht und Roswitha Văth-Szusdziara, 1976: Gesamtschule und dreigliedriges Schulsystem. Eine Vergleichsstudie über Chancenungleichheit und Durchlässigkeit. Stuttgart: Klett.
- Helberger, Christof, 1980: Veränderungen der bildungsspezifischen Einkommensunterschiede zwischen 1969/71 und 1978. Arbeitspapier Nr.51 des Sonderforschungsbereich 3 der Universität Frankfurt. Frankfurt.
- Kyn, Oldrich, 1978: Education, Sex and Income Inequality in Soviet-type Socialism. S.274-289 in: Zwi Griliches, Wilhelm Krelle, Hans-Jürgen Krupp und Oldrich Kyn (Hrsg.), Income Distribution and Economic Inequality. Frankfurt: Campus.
- Pettmann, Barrie O. (Hrsg.), 1975: Equal Pay for Women. Progress and Problems in Seven Countries. Bradford.
- Schmidt, Folke (Hrsg.), 1978: Discrimination in Employment. A Study of Six Countries by the Comparative Law Group. Stockholm.
- Statistisches Bundesamt, 1970-1980: Tarifgehälter. Fachserie 16, Reihe 2.2. Stuttgart, Mainz: Kohlhammer.
- Swafford, Michael, 1978: Sex Differences in Soviet Earnings. *American Sociological Review*, 43: 657-673.
- Treiman, Donald J., 1979: Begriff und Messung des Berufsprestiges in der international vergleichenden Mobilitätsforschung. S.124-167 in: Franz Urban Pappi (Hrsg.), Sozialstrukturanalysen mit Umfragedaten. Königstein/Ts.: Athenäum.
- Winsborough, H.H. und Peter Dickinson, 1971: Components of Negro-White Income Differences. *Proceedings of the American Statistical Association*: 6-8.

Die Autoren dieses Bandes

Andreas Diekmann, geb. 1952, Dipl.-Soz., Dr. rer. pol., Assistent am Institut für Höhere Studien, Wien

Hartmut Esser, geb. 1943, Dipl.-Soz., Dr.rer.pol., Professor für empirische Sozialforschung an der Universität Essen

Jürgen H.P. Hoffmeyer-Zlotnik, geb. 1946, Dr. phil., Projektleiter bei ZUMA

Hans-Peter Kirschner, geb. 1945, Dipl.-Math., Dr. rer. nat., Leiter der Statistikabteilung bei ZUMA

Cornelia Krauth, geb. 1954, Dipl.-Soz., wiss. Mitarbeiterin des Projekts ALLBUS, Universität Mannheim und ZUMA

Manfred Küchler, geb. 1943, Dipl.-Math., Dr. rer. nat., Professor für Statistik in den Sozialwissenschaften, wissenschaftlicher Leiter bei ZUMA

Karl Ulrich Mayer, geb. 1945, M.A., Dr. rer. soc., Professor für Soziologie, Direktor am Max-Planck-Institut für Bildungsforschung in Berlin, bis 1983 wissenschaftlicher Leiter des Projekts ALLBUS

Rolf Porst, geb. 1952, Dipl.-Soz., wiss. Mitarbeiter des Projekts ALLBUS, Universität Mannheim und ZUMA

Volker Schanz, geb. 1949, Dipl.-Soz., wiss. Mitarbeiter im Sonderforschungsbereich 24 ("Sozial- und wirtschaftspsychologische Entscheidungsforschung"), Universität Mannheim

Peter Schmidt, geb. 1942, Dipl.-Soz., Dr. phil., Professor für empirische Sozialforschung an der Universität Gießen

Gunter Wolf, geb. 1948, Dipl.-Vw., wiss. Mitarbeiter des Projekts "Westdeutsche Elite 1981", Universität Mannheim

ISBN 3-593-33262-0